



HAL
open science

Contribution à la cartographie d'une matrice de flux

Françoise Bahoken

► **To cite this version:**

Françoise Bahoken. Contribution à la cartographie d'une matrice de flux. Géographie. Université Paris Diderot (Paris 7), 2016. Français. NNT: . tel-01273776

HAL Id: tel-01273776

<https://shs.hal.science/tel-01273776>

Submitted on 13 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

présentée pour obtenir le grade de

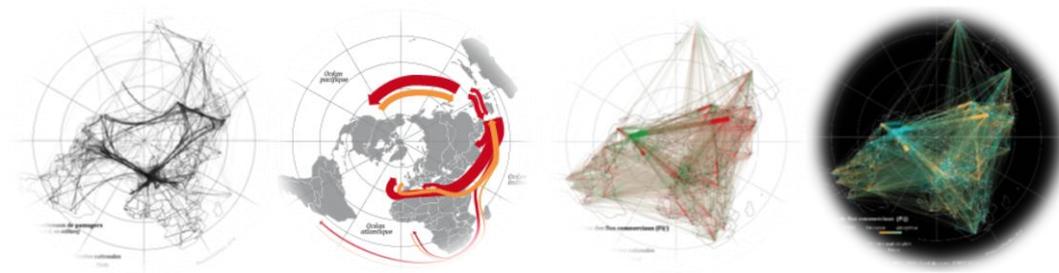
Docteur de l'Université Paris Diderot (Paris 7)

Discipline : Géographie - Sciences des territoires

Ecole Doctorale : Géographie de Paris (ED 434). Espaces, Sociétés Aménagement

Françoise BAHOKEN

Contribution à la cartographie d'une matrice de flux



réalisée sous la direction de :

Claude GRASLAND et de Christine ZANIN

*Soutenue publiquement le 5 janvier 2016,
devant le jury composé de :*

Madame Anne BRETAGNOLLE	Professeur à l'Université Panthéon-Sorbonne, Paris 1. <i>Présidente.</i>
Monsieur Denis ECKERT	Directeur de Recherches au CNRS, Toulouse. <i>Rapporteur.</i>
Monsieur Claude GRASLAND	Professeur à l'Université Paris Diderot, Paris 7. <i>Directeur.</i>
Monsieur Alain L'HOSTIS	Chargé de Recherches - HDR à l'IFSTTAR, Lille. <i>Rapporteur.</i>
Madame Paule-Annick DAVOINE	Maître de conférences - HDR à Polytechnique, Grenoble, <i>Examinatrice.</i>
Madame Christine ZANIN	Maître de conférences - HDR à l'Univ. Paris Diderot, Paris 7. <i>Co-directrice</i>

AVANT PROPOS

Cette thèse a été réalisée à l'Université Denis Diderot Paris 7, au sein de l'UMR n° 8504 GEOGRAPHIE-CITES (CNRS-Université Paris 1 et Paris 7) - Equipe *Pour l'avancement des recherches sur l'interaction spatiale* (PARIS), à la marge d'une activité professionnelle principale à temps plein d'Ingénieur d'études en *Traitement et représentation de l'information statistique et spatiale* (titre exact du métier de cartographe dans le secteur de l'enseignement supérieur et de la recherche), et d'activités secondaires (Chargée de cours et de Travaux dirigés et/ou de tutrice, de « bidouilleuse informatique »¹...).

Ma réflexion a ainsi pu être enrichie par mes activités professionnelles quotidiennes et par leur environnement relationnel : d'abord au sein de l'UMR n°6588 MIGRINTER (CNRS et Université de Poitiers), spécialisée dans l'étude des migrations internationales de populations puis à l'unité de recherche SPLOTT de l'IFSTTAR (ex-INRETS) dédiée notamment à la logistique du transport de marchandises. Aussi par ma participation nécessairement irrégulière aux groupes de travail *Cartomouv'* (*Cartographie du mouvement, Cartographie du Changement*), au groupe *fnr* (*flux, matrices, réseaux*), au destin éphémère (?) et, plus récemment à *ElementR*. Enfin et c'est non le moindre, par le réseau de relations de GEOGRAPHIE-CITES, en particulier avec l'UMS RIATE et l'UMR IDEES, pour ce citer que ces deux structures.

Si cette thèse semble avoir couvert un pan important de mon existence professionnelle, ce qui est réellement le cas, il m'apparaît important de préciser que ma recherche a fait l'objet de deux sujets aujourd'hui bien distincts.

Le premier d'entre eux, intitulé *Cartographie des flux de population : méthodes et concepts*, était focalisé sur la représentation graphique des déplacements de populations humaines sous un double aspect individuel et agrégé. Il n'a pu aboutir au terme de trois années de travaux pour différentes raisons qui peuvent toutefois être rassemblées sous un motif commun : l'extension trop importante du domaine de la recherche. L'objectif initial d'articuler notions et concepts qui relèvent d'acceptions variées des mobilités spatiales s'est en pratique avéré impossible à tenir. D'où la notification, en 2011, par le comité de thèse constitué à l'époque, d'un « *abandon administratif* » du sujet.

Le second sujet est déposé pour l'année universitaire 2013-2014. S'il est issu du premier, il en diffère toutefois du point de vue de l'objet, en étant désormais focalisé sur l'analyse d'une matrice de flux unique dans le cadre d'une démarche théorique et méthodologique d'ensemble plutôt que sur les seules représentations graphiques. Si les questions liées au cheminement individuel sont abandonnées au profit de l'agrégat, nous avons également souhaité étendre le propos à différents types de flux (de population, de marchandises, de valeurs, ...) pour assurer *a priori* le caractère général et généralisable de la démarche proposée. Cela étant, ma sensibilité personnelle à la thématique des migrants et autres migrations transparaîtra fatalement dans le texte, étant donnée la difficulté de tout recommencer, fusse la direction de la thèse, vraiment, excellente.

¹ Qualificatif qui me fut attribué publiquement, au début de ma carrière professionnelle dans le secteur public.

REMERCIEMENTS

La page la plus difficile à écrire est incontestablement celle des remerciements, car de nombreuses personnes ont permis la réalisation de ma thèse, puis sa finalisation. Au-delà des individus, j'accorde une importance au contexte : aux conditions et au ressenti des situations, aux moments privilégiés, aux regards ; aussi aux questions, aux commentaires, remarques et suggestions. Tout cela participe de ces *interactions* sociales qui ont indéniablement enrichi mon travail, participé à ma réflexion. Dans l'ensemble des personnes que j'ai pu croiser dans le cadre de ma thèse, il en est que je tiens particulièrement à remercier, non seulement parce qu'elles ont une position *dominante* dans mes recherches, ce qui est normal, mais aussi parce qu'elles m'ont marquées par leur(s) attention(s) bienveillante(s). L'exercice n'est plus seulement difficile. Il est désormais périlleux, car il va me conduire à ne mentionner qu'une *sélection* de *relations* répondant à ces *critères*, celles que l'on va alors considérer comme *significatives*, ce qui est pour le moins injuste pour les personnes que j'aurais omises. En effet, ce n'est pas tant la quantité d'interactions que leur qualité qui m'a touchée. C'est pourquoi je tiens à m'excuser par avance auprès des personnes que je n'aurai pas mentionnées à tort et leur demande de ne pas m'en tenir rigueur.

J'ai choisi Claude Grasland pour diriger mes recherches et lui adresse aujourd'hui le *flux majeur* de mes remerciements, pour avoir accepté de me suivre pendant toutes ces années et m'avoir permis de mener ce projet à son terme. Je lui suis également très reconnaissante d'avoir d'une part, positionné ma *trajectoire* (dès le 1^{er} jour de notre rencontre) sur les traces de Waldo Tobler qui m'ont ensuite permis de trouver ma propre voie. D'autre part, d'avoir rapidement associé l'approche « rétinienne » de Jacques Bertin et par voie de conséquences, Christine Zanin, à la direction de ma thèse. Je souhaite la remercier dans ce même élan de reconnaissance pour toutes nos discussions, pour son attention et pour son attentive oreille. Je dois dire que Claude Grasland et Christine Zanin ont constitué le *couple* de directeurs (pas de lieux, bien sûr !) idéal, le binôme affûté d'une thèse mobilisant simultanément aspects statistiques et considérations cartographiques. La confrontation de leurs visions (pas toujours convergentes) nous a d'ailleurs valu des réunions de travail épiques, mais toujours stimulantes, en tous les cas très enrichissantes. Un grand merci à vous deux, ce fut un plaisir.

J'adresse mes remerciements respectueux aux membres de mon jury : à mes deux rapporteurs Denis Eckert et Alain l'Hostis que j'ai beaucoup cité ; à mes deux examinatrices : Anne Bretagnolle, par ailleurs présidente et Paule-Annick Davoine, dont la présence me renvoie au début de ma formation universitaire, à Grenoble.

Cette recherche n'aurait pu aboutir sans la compréhension inégalée de Corinne Blanquart, directrice de mon laboratoire d'affectation (AME - SPLOTT) à l'IFSTTAR. Je souhaite lui exprimer ma reconnaissance infinie, ne sachant comment j'aurai pu la remercier autrement.

Corinne fait partie de celles et de ceux dont la présence furtive, *discrète* ou en *continu* ..., plus généralement des personnes dont l'attitude bienveillante fut une source inestimable d'encouragements. C'est pourquoi je tiens à remercier en particulier, par ordre alphabétique : Arnaud Banos, Myriam Baron, William Berthomière, Sophie Beudet-Michel, Clarisse Didelon, César Ducruet, Grégoire Feyt, Antoine Fleury, Antoine Frémont, France Guérin, Olivier Iyebi-Mandjek, Laurent Jégou, Lilianne

Lizzi, Malika Madelin, H  l  ne Mathian, Marie Morelle, Na  k Miret, Patrick Ni  rat, Denise Pumain, Anne Ruas, Sophie de Ruffray, L  na Sanders (et les trois inconnu.e.s ...), Camille Schmoll, Muriel Sam  -Ekobo et j'en oublie... Je ne saurai toutefois omettre de mentionner   galement Martine Laborde, Saber Marrouchi et V  ronique Degout ; Jocelyne Cabald ; Maurad Hama  idi ; Nicole Verdi  re.

I also want to thank a foreign personality for his interest and above all documents about mapping migration flows patterns he sent to me: Professor David A. Plane from the University of Arizona.

J'ai aussi la chance d'avoir pu collaborer avec des coll  gues moins   loign  s, situ  s dans mon *voisinage* professionnel, ce qui me conduit    saluer vivement mes co-auteurs. D'abord Ana-Maria Olteanu-Raimond pour touuuuuus nos travaux pass  s et    venir, en particulier pour la retomb  e de l'ACI de Dresdes (d'ailleurs gr  ce    Anne Ruas que nous remercions, encore, pour son support)    savoir ces messages d'encouragements re  us de Waldo Tobler, en personne. Ensuite Antoine Fr  mont pour tous ces flux, non : pour tous ces *mouvements* par voie maritime r  alis  s    travers le monde. Olivier Bonin et Samuel Mermet pour nos   pop  es autour de CARTOD et ce prix du FIG ! A ce propos, j'en profite pour saluer Isabelle Bouhet - Andr  -Poyaud dont la collaboration nous a valu une pr  c  dente distinction au FIG. Mes camarades du *groupe Cartomouv'* ainsi que ceux du *groupe fmr*, Laurent Beauguitte et Serge l'Homme pour l'ECTQG de Dourdan, Matthieu Drevelle et Marion Maisonobe pour Netcom :    tous, merci pour nos discussions, s  minaires et autres JE. Et hop, avec un *lien* plus loin, je me retrouve    *ElementR* : j'y salue amicalement ses participants et remercie en particulier Paul Chapron, pour la s  ance 5. Avant ceux la, il y a eu l'axe 5 transversal de Migrinter que je n'ai pas oubli  , aussi Christophe Imbert, Yann Scioldo-Z  rcher (pour les p  r  grinations vers le Canada), Olivier Clochard pour les atlas des migrants en Europe, et Gilles Dubus pour la bibliographie sur les migrations, pour e-migrinter, ...

Je remercie pour leur collaboration chaleureuse    la relecture, en fonction de leurs sp  cialit  s : Laurent J  gou notamment pour la *Couleur*, facile, mais pas seulement ; Clarisse Didelon, Muriel Sam  -Ekobo et Marie Morelle notamment pour la com-pr  -hen-sion ; Lilianne Lizzi, Emilie Gaubert et Antoine Montenon pour les traitements ; Laurent Beauguitte pour la section sur les graphes ; Genevi  ve Maurel pour l'orthographe, Sylvain Cuyala pour l'imprimeur. Pour ma soutenance, je souhaite remercier Charl  ne le Neindre, C  sar Ducruet, Nicolas Lambert et Ronan Ysebaert.

Dans ma th  se, l'  quipe de direction a aussi compt   par sa pr  sence en *continu*. Je m'en souviendrai toujours comme de pr  c  dents professeurs qui ont marqu   mon apprentissage, parmi lesquels par ordre chronologique : Monsieur Mbida Ab  da, Mesdames Kalliopi Ango-Ela et Monique Mallo (   Yaound  , Cameroun) ; Monsieur Dolmagian (   La Mure, France) ; Monsieur Fr  d  ric Giraut pour l'id  e du classeur de cartographie statistique que j'ai encore. Et, *last but not least* : Monsieur Pierre Dumolard, pour m'avoir donn   le « gout du gombo » en dirigeant ma ma  trise et mon DEA (   Grenoble).

J'ai une pens  e particuli  re pour Henri Chamussy, un ancien professeur, Herv   Rakoto Ramiarantsoa et Matthieu Giroud, d'anciens coll  gues qui connaissaient bien mon projet mais ne sauront jamais, h  las, qu'il a pu aboutir.

J'adresse enfin un clin d'oeil    mes proches pour leur accompagnement et leur soutien ind  fectible pendant toutes ces ann  es : Oliv', Man's, Mur, Flo, Caro, Laura, Marie-H  l  ne, Marie, Carine, ...

A Sylvain, mon compagnon que je remercie affectueusement pour tous nos moments de vie et le support pour l'  criture des programmes Rstudio annex  s. A H  lo  se et Agathe, nos jujues, pour tous leurs dessins de « la terre des flux » et surtout pour avoir bien compris, malgr   leur jeune   ge, l'importance que rev  tait pour moi la finalisation de cette th  se.

A Marion.

RESUME

Une matrice de flux est un tableau croisé formé de $(1, \dots, i, \dots, k)$ lignes et de $(1, \dots, j, \dots, p)$ colonnes, respectivement lieux d'origine et de destination, dont le croisement forme une valeur quantitative : le flux (F_{ij}) . Sa cartographie a pour objectif de représenter graphiquement ces valeurs. Pour cela, elle les transforme en figurés linéaires, ponctuels et/ou zonaux puis les projette dans un espace assorti d'une métrique communément géographique. Lorsque la matrice est dense, cette approche pourtant bien connue pose de nombreuses difficultés d'ordres méthodologique et pratique, des contraintes graphiques et des défis techniques et technologiques qu'il convient de résoudre au risque de produire une figure illisible, caractéristique de *l'effet spaghetti*. La réponse traditionnellement apportée consiste à simplifier la figure en sélectionnant l'information à représenter et/ou en réduisant sa résolution. Essentiellement numérique ou technologique, elle ne s'inscrit pas dans un raisonnement théorique et conceptuel qui intéresse le processus de construction cartographique. Quelles en sont les raisons ? Est-ce à dire que les solutions à ces problèmes cartographiques seraient exclusivement numériques ? Que signifie *in fine* représenter des valeurs de flux sur une carte ?

La thèse s'inscrit dans l'objectif général de réduction de la complexité inhérente au processus de cartographie d'une matrice de flux. Elle propose pour ce faire quatre familles de contributions théoriques et méthodologiques. La première d'entre elles est d'ordre analytique, elle fournit une formalisation des composantes de *l'effet spaghetti* conduisant à une typologie (validée sur des cartes anciennes et récentes) associant une et une seule méthode de cartographie aux trois principaux types de matrices de flux identifiés. La seconde contribution est théorique et conceptuelle. Fondée sur une grille d'analyse sémiotique (signifiant, signifié), elle établit la signification de la carte de flux au regard de celle du mouvement et introduit l'espace géographique dans le processus cartographique (en complément éventuel de sa modélisation qui ne fait pas l'objet de la thèse). En montrant l'invariance du processus cartographique du flux, quelle que soit l'échelle géographique et la thématique, la thèse souligne l'importance de la sémantique (flux, mouvement) et, par conséquent, de l'interprétation et de la procédure sous-jacente de sélection de l'information de flux représentée. La troisième famille de contribution est d'ordre méthodologique. Elle propose, dans un premier temps, un renouvellement de la sémiologie cartographique des flux qui modifie la perception de ses motifs sans mobiliser les données numériques. La quatrième famille de contribution porte le regard, à l'inverse de la précédente, sur les données numériques, en particulier sur le choix du critère de sélection de l'information à représenter. D'abord fondée sur l'intégration de différents critères spatiaux de séparation ou de proximité des lieux pour réduire la matrice, l'approche mobilise ensuite un critère unique de sélection qui agit soit globalement, c'est-à-dire à l'échelle de la matrice, soit localement, autrement dit sur les localités (les lignes et/ou les colonnes) introduites dans le cadre d'une analyse cartographique des flux dominants. L'approche se voulant générale et généralisable, les solutions proposées sont validées empiriquement sur des matrices de flux migratoires ou commerciaux s'exprimant à différentes échelles géographiques (de l'échelle locale au niveau mondial).

Mots-clés : matrice de flux, carte de flux, carte de mouvement, flux dominants, figuré, procédé, distance cartographique, voisinage, similarité, perception, cognition, Bertin, Tobler, Nyusten et Dacey. Rstudio©.

ABSTRACT

We consider a flow matrix where the $(1, \dots, i, \dots, k)$ rows and the $(1, \dots, j, \dots, p)$ columns are respectively origin and destination place linked by a flow value (F_{ij}) . Mapping flow aims to project values on a map depicting generally a geographic space, using linear and punctual features. When the matrix is dense, this well known approach leads theoretical and methodological difficulties, graphical constraints, technical and technological challenges that must be addressed in order to solve the graphic complexity that appears: the so-called *spaghetti effect*. The traditional answer is to simplify the figure by selecting flows and or by reducing their resolution by clustering procedures. If the result is really effective, it does not interest the flow mapping cartographic process. What are the reasons? Does this mean that solutions to flow mapping problems are only related to digital or cartographic data? What is the real meaning of flow mapping?

This PhD is part of the overall objective of reducing analytic and graphical complexity of flow mapping. We offer here four families of contributions. First of all, an analytical contribution that provides a formalization of the components of the *spaghetti effect*. That one leads to a typology of flow mapping methods (validated on both old and recent maps) that combines graphical feature and main type of flow matrix. Our second family of contributions is theoretical and conceptual. Based on a semiotic analysis grid (about the signifier and the signified information on flows), the thesis establishes the meaning of the flowmap compared to the movement map by introducing the geographical space in the mapping process. By showing the geographic and thematic invariance of the general process of mapping flows (or movement), it highlights the importance of the semantic component of flow mapping, then the interpretation and, finally, the information selection process. The third and fourth families of contributions are methodological. The thesis offers, in the third one a renewal of the flowmap semiotics that changes the features' information perception without mobilizing flow data and, in the fourth', several ways of selecting flow values to map. The criterion is firstly based on exogenous data: on places' separation. The approach mobilizes secondly a single criterion which is either globally (to apply across the whole matrix), either locally, that is to say focused on places introduced as part of a mapping approach of nodal flows mapping. The general approach is wanting general and generalizable, the solutions we offer are empirically validated on different matrices (migration, trade, financial flows,...) dealing at several geographical scales, from local to global levels.

Key-words: flow matrix, flow mapping, movement mapping, nodal flows, graphic feature, cartographic process, cartographic distance, neighborhood, similarity, perception, cognition, Bertin, Tobler, Nyustén and Dacey, Rstudio(c).

SOMMAIRE

PREMIERE PARTIE ETAT DE L'ART DES MÉTHODES DE CARTOGRAPHIE DE FLUX. ÉNONCÉ DES PROBLÈMES, DES CONTRAINTES GRAPHIQUES ET DES DIFFICULTÉS D'ORDRE MÉTHODOLOGIQUE	25
1 ^{ER} CHAPITRE. LA CONSTRUCTION DE LA MATRICE DE FLUX : INCERTITUDES DE MESURE ET PROBLÈMES D'AGRÉGATION	27
2 ^{ÈME} CHAPITRE. LES MODÈLES DE LA REPRÉSENTATION GRAPHIQUE DES MATRICES DE FLUX : LA MATRICE ORDONNÉE, LE GRAPHE ET LA CARTE	55
3 ^{IÈME} CHAPITRE. LA DÉCOMPOSITION DES MATRICES DE FLUX (FIJ) : LE VOLUME (FIJ ⁺) ET LE SOLDE (FIJ ⁻) BILATÉRAL	109
4 ^{IÈME} CHAPITRE. L'ÉVOLUTION DES MÉTHODES DE CARTOGRAPHIE DE FLUX : ÉTAT DE L'ART DES PROCÉDÉS ET DES PRINCIPES DE SÉMILOGIE GRAPHIQUE	141
DEUXIEME PARTIE PROPOSITIONS DE SOLUTIONS VISANT À RÉDUIRE LA COMPLEXITÉ DE LA CARTE DE FLUX.....	205
5 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES AUX ASPECTS SÉMILOGIQUES ET ESTHÉTIQUES DE LA CARTE DE FLUX : LE RENOUVELLEMENT DE LA VARIABLE VISUELLE <i>COULEUR</i>	211
6 ^{IÈME} CHAPITRE. D'UNE CARTOGRAPHIE DE FLUX (FIJ) À UNE CARTOGRAPHIE DE MOUVEMENTS (MIJ) ET INVERSEMENT	247
7 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES À LA PRISE EN COMPTE DE L'ESPACE : LA DISTANCE, LE VOISINAGE SPATIAL ET L'APPARTENANCE TERRITORIALE	285
8 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES AU CHOIX D'UN CRITÈRE DE SÉLECTION GLOBAL : LA CARTOGRAPHIE DE <i>FLUX SIGNIFICATIFS</i> (FIJ> α)	325
9 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES AU CHOIX D'UN CRITÈRE DE SÉLECTION LOCAL : LA CARTOGRAPHIE DE <i>FLUX DOMINANTS</i>	347
BIBLIOGRAPHIE	1
LISTE DES TABLES.....	25
GLOSSAIRE.....	33
NOTATIONS.....	41
TABLE DES MATIERES.....	43
ANNEXES	1

INTRODUCTION GENERALE

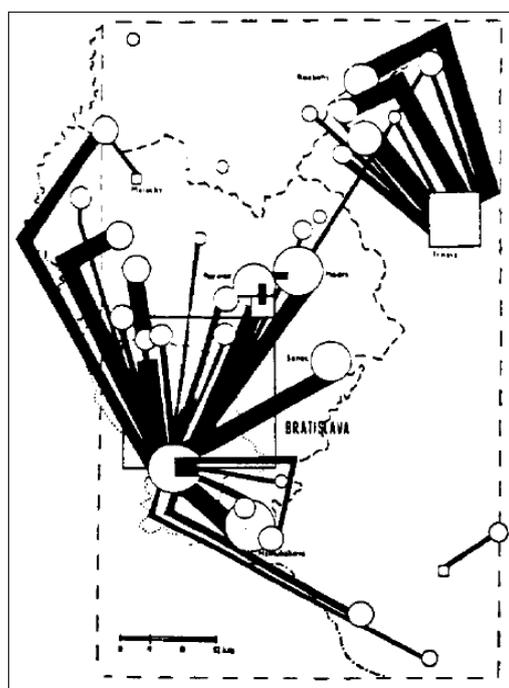
« La cartographie n'est pas seulement une représentation visualisée des localisations. C'est aussi l'expression synthétique d'une masse d'information dont le traitement et la traduction graphique exigent une recherche méthodologique d'ensemble ».

Peter Haggett (1973)

Il est aujourd'hui acquis que la cartographie des matrices de flux pose problème, dès lors que la zone d'étude comporte de nombreux lieux, en tous les cas plus d'une dizaine. En effet, l'image qui en résulte se caractérise par une complexité graphique intrinsèque la rendant « [...] illisible et donc inutile » (Torrice, 1997 : 9). Pour y remédier, un certain nombre de solutions diverses sont déployées pour réaliser une figure intelligible, des cartes suggestives à fort impact visuel généralement d'une incroyable clarté, d'une « netteté » que W. Tobler (n.d-1) appréhende d'ailleurs comme une propriété générale de la carte de flux.

Effectivement, la représentation cartographique typique d'une matrice de flux peut être illustrée par la Figure I-1, proposée par R. Brunet (1987), en page 89 de *La Carte mode d'emploi*, pour accompagner la définition des cartes sagittales.

Figure I-1-1. Une carte de flux



Loisirs en Slovaquie.

« Carte des déplacements de loisirs autour de Bratislava et Trnava : vers les collines des Petites Carpates et sur les bords du Danube – en arrière de la frontière (Hamuliakovo). Bandes de largeurs, proportionnelles aux flux.

Par P. Mariot, *Geographicky Casopis* 1976 n°1 ».

Source : R. Brunet (Figure 29.1)

La méthode adoptée pour représenter les flux de la Figure I.1 est mentionnée dans la légende (placée ici sur la droite), par l'inscription : « Bandes de largeurs, proportionnelles aux flux ».

L'auteur de cette carte, P. Mariot, a-t-il raison de représenter ces déplacements de loisirs en Slovaquie comme il le fait, c'est-à-dire à l'aide de bandes – « *de largeurs proportionnelles, aux flux* » – de cercles et de carrés ?

Cette question nous permet d'introduire le propos de cette thèse à savoir contribuer à la représentation graphique des matrices de flux denses, sous la forme d'une carte statique² (telle que présentée sur la Figure I.1), quel que soit le support (papier ou numérique). Pour ce faire, nous allons nous intéresser aux aspects d'ordre méthodologique qui concourent à sa réalisation, en focalisant notre attention sur un double choix : celui de l'information de flux à représenter (celle qui est disponible dans le tableau) et celui des méthodes de représentation qui lui sont associées (une bande, par exemple), en tenant compte des aspects liés à l'interprétation de l'image des flux qui en résulte, c'est-à-dire à la perception des motifs de la carte par un observateur. L'une des questions majeures qui se pose est incontestablement celle de la compréhension des motifs de flux représentés par l'enchevêtrement de bandes représentées sur une carte de flux, sur la concomitance des figurés linéaires et ponctuels, voire zonaux.

En géographie, plus largement en sciences humaines et sociales, la cartographie de flux mobilise des données quantitatives. C'est pourquoi la carte qui en résulte est dite thématique ou statistique. Elle représente, d'après le dictionnaire *Les mots de la Géographie*, « [...] *la distribution spatiale d'un phénomène, à partir de relevés statistiques, de données recueillies par satellites, ou d'autres sources.* » (Brunet et al., 1992 : 89). Au même titre que d'autres types de cartes, la carte thématique décrit des « [...] *phénomènes invisibles, non susceptibles d'être délimités sur le terrain (densités de populations, revenus, pluies), mais mesurés ou évalués* ». Parmi la typologie proposée par l'auteur, la cartographie thématique d'une matrice de flux correspond au type « [...] *7. cartes sagittales : elles montrent des trajectoires, avec des lignes plus ou moins épaisses et colorées représentant des liaisons (réseaux) ou des flux (échanges) entre des points ou des circonscriptions* » (Brunet et al., 1992 : 89-90).

La carte de flux qui nous intéresse est exclusivement réalisée selon un mode d'implantation spatiale par définition discret : le fond de carte est formé de points et/ou de polygones qui représentent des lieux ; l'information représentée est symbolisée par des formes géométriques, comme sur la Figure I-1. Les lignes sont, par exemple, associées à des flux. Nous n'allons pas examiner le cas des cartes de flux réalisées sur des surfaces continues ou de tendance non adaptées à tous les types de flux.

Les cartes de flux sont issues de tableaux spécifiques appelés matrices de flux, un objet mathématique présentant de nombreuses propriétés analytiques et correspondant à différentes formalisations graphiques. Elles sont liées au type de flux (planaires / non planaires ; matériels / immatériels ; inter zones / intra zones, de populations / de biens, ...) dont les caractéristiques influent nécessairement sur le résultat cartographique. Aussi allons-nous réduire d'emblée notre propos aux seuls flux dits Origine-Destination (*flux OD*)³.

Nous définissons en première instance le *flux* comme le transfert direct, c'est-à-dire en ligne droite, d'une quantité entre un lieu d'origine et un lieu de destination. Ce flux n'est pas *affecté sur réseau*, ce qui signifie d'une part, que l'on ne tient pas compte du mode de transport emprunté entre l'origine et la destination et d'autre part, que son cheminement exact n'est pas connu. Nous justifions cette

² La généralisation des supports-écran a naturellement conduit à l'émergence de réflexions méthodologiques sur la représentation animée des mouvements et des flux. Ces approches sont principalement centrées sur les aspects de sémiologie graphique (Cheylan, 2007, Kaddouri, 2008, Zanin et Ben Rebah, 2011) qui concourent à la définition de variables visuelles de l'animation plus aptes à représenter le mouvement.

³ Les termes et expressions portés en italiques sont explicités dans le texte, puis récapitulés sous une forme synthétique dans le Glossaire, situé à la fin du document.

restriction aux seuls *flux OD* en aspect direct (prenant toutefois en compte en compte le cheminement dans les derniers chapitres) par des raisons d'ordre exclusivement méthodologique. Les flux *affectés sur réseaux* présentent la particularité d'être liés à une infrastructure matérielle : un réseau planaire, généralement formé de sous-composantes connexes qu'il est nécessaire d'intégrer dans l'analyse. Cela suppose le recours à des méthodes spécifiques, telles que, la modélisation du trafic du point de vue de la demande ou de l'offre de transports, ou encore l'analyse des différentes possibilités de cheminements dans le graphe correspondant au réseau de transport. Si la cartographie de flux affectés sur réseaux peut être considérée comme un cas particulier de la cartographie des *flux OD* et inversement (la cartographie de flux comme une variante de celle du réseau), elle met en œuvre des notions qui ne sont pas toujours compatibles entre elles, car elles mobilisent des concepts différents. En effet, un flux sur réseau est relatif au cheminement spatio-temporel d'une quantité sur une infrastructure de transport (routière, ferroviaire,...) : c'est un *flux matériel*. A l'inverse, un *flux OD* considère le transfert direct d'un agrégat entre des OD, dans un espace quelconque : il est caractéristique d'un *flux immatériel*.

Genèse et problématisation du sujet

La carte de flux est un objet peu représenté en géographie. Néanmoins, elle suscite un intérêt marqué pour un public hétéroclite : géographes, démographes, économistes mais aussi artistes, graphistes et, dans des domaines plus techniques, informaticiens, statisticiens, mathématiciens, modélisateurs, ... ; les enseignants et les étudiants, de l'université à l'école primaire, s'y intéressent notablement. Plusieurs raisons peuvent être avancées, notamment la simplicité apparente de la figure. La carte de flux – lorsqu'elle est bien faite – est formée d'une sélection d'objets graphiques élémentaires : des points et des lignes, souvent disposés de manière éparse sur un fond de carte, dans un souci évident d'harmonie et de clarté visuelle. Cette simplicité apparente cache pourtant une extrême complexité analytique et graphique.

La carte de flux est le résultat de l'articulation graphique de plusieurs dimensions (thématiques, temporelles et spatiales) d'une même matrice de flux, qui en font un objet complexe à élaborer, à analyser. Cette complexité analytique engendre une complexité graphique qui n'est que rarement évoquée. En effet, leur représentation graphique pose problème, en particulier pour les grandes matrices, celles qui sont formées de plus d'une dizaine de lieux. « *La représentation des migrations régionales. C'est un problème classique, très complexe. En effet : 1°) Les mouvements observés peuvent relier toute zone administrative à toute autre zone, et ceci dans les deux sens. Le seul réseau des mouvements orientés est donc très complexe.* » (Bertin, 1973 : 351). La réponse traditionnellement apportée à ce problème consiste à sélectionner *a priori* les données sans se poser de questions sur la manière dont l'information est sélectionnée. Pour preuve, la carte de R. Brunet (voir figure I.1) ne représente qu'une sélection de flux entre une sélection de lieux, ceux jugés comme étant les plus importants par l'auteur. Cette opération de sélection, correspondant à une réduction de la quantité d'information disponible, est-elle obligatoire ? Que traduit-elle ? N'y aurait-il pas d'autre manière de procéder ? Comment l'information représentée a-t-elle été sélectionnée (par quelles méthodes) ? Quels sont les (types) de critères mobilisés et ceux qui pourraient l'être ?

On pourrait penser que l'information de flux est sélectionnée ou non en fonction de la thématique du flux et, également, en fonction des choix réalisés par l'auteur. Généralement, il s'agit de décider d'un seuil de valeurs à partir duquel les flux seront représentés (ou éliminés). Cependant, procéder ainsi – choisir empiriquement l'information de flux à représenter – n'assure ni la robustesse ni l'exactitude du résultat obtenu : « [...] bien que les ordinateurs aient automatisé le mécanisme de production d'un graphe, ils n'ont pas automatisé la tâche de choisir le bon type de graphe qui illustre le mieux l'information [...] alors qu'il est facile d'inventer un nouveau type de graphe, il est difficile d'en

inventer un qui marche bien » (Lohse, 1997, cité par Otjacques, 2004 : 23). Hormis le logiciel Flowmap⁴, dont le manuel d'utilisation consacre deux pages à la sélection⁵ des données de flux (chapitre 4), il n'existe pas, à notre connaissance, de documentation logicielle (en activité⁶) décrivant explicitement la procédure de sélection des flux à représenter sur une carte.

Il est pourtant nécessaire, pour réaliser une sélection de flux à cartographier, de combiner différentes procédures de sélection de l'information qui exigent une bonne connaissance du fonctionnement des matrices, dans le contexte de l'analyse cartographique. En d'autres termes, une bonne connaissance du maniement des matrices d'information géographique (qui décrivent la zone d'étude) et de celles de flux est indispensable ; en même temps qu'une bonne connaissance de la thématique des flux est nécessaire. Etant donné que l'information sous-jacente à la réalisation d'une carte de flux est de nature statistique, nous formulons l'hypothèse **qu'il est possible de disposer de méthodes solides de sélection de l'information de flux à cartographier. Leur utilisation présenterait l'avantage d'assurer la robustesse du résultat obtenu, ainsi que le caractère généralisable et reproductible de leur construction.**

R. Brunet précise, dans sa description des cartes de flux et de réseaux, l'importance de la formalisation graphique géométrique des cartes de flux, par rapport à celles des réseaux, pour résoudre un problème de lisibilité⁷ spécifique à la carte de flux. « *Pour être lisibles, les cartes de flux sont souvent traitées comme des graphes : les mouvements d'échanges sont représentés sans tenir compte des itinéraires réels, selon des modèles géométriques* » (Brunet, 1987 : 90). Cette phrase soulève la question de la différence entre, d'une part, la carte de flux et le graphe de flux (ou le graphe valué) et, d'autre part, entre la carte de flux et celle, encore plus rare, de « *mouvements d'échanges* », c'est-à-dire de flux affectés sur réseau ou sur l'itinéraire le plus probable. Cette assertion de R. Brunet (*ibid.*) nous conduit à plusieurs questions de fond quant à la signification d'une carte de flux et à sa représentation. Est-ce réellement la carte de flux qui est modélisée, dans sa représentation, comme un graphe ou ces deux objets sont-ils formellement identiques ?

La carte de flux, tout comme le graphe, correspond effectivement à la projection dans un espace géographique ou géométrique des valeurs d'une matrice, ce qui nous conduit à poser l'hypothèse de **la similarité du graphe et de la carte de flux au niveau de leur construction**. Nous posons le fait que le graphe, la carte de flux et celle de réseaux correspondent en fait, tous trois, à la représentation sous la forme de points et de lignes de valeurs disponibles dans une matrice. Partant de là, nous pouvons alors nous poser la question de la similarité de la construction graphique de ces trois objets.

La seconde question soulevée par R. Brunet (*ibid.*), que nous avons d'ailleurs très souvent rencontrée dans la littérature, est d'ordre théorique. Elle concerne l'utilisation de plusieurs dénominations pour qualifier, apparemment, un même objet. Les expressions suivantes semblent effectivement substituables l'une à l'autre : « *cartes de flux* » (« *flow map* ») (Pumain, Saint-Julien, 2005 ; Dorling,

⁴ Flowmap© est l'unique Système d'Information Géographique (SIG) entièrement dédié aux flux. Développé à l'Université d'Utrecht, il présente une version « étudiants » quasi complète et disponible librement en ligne à l'adresse suivante : <http://flowmap.geo.uu.nl/> [Lien vérifié le 02/08/2015], la version professionnelle pouvant être acquise pour une somme modique.

⁵ Ce quatrième chapitre du manuel de Flowmap© est composé de deux sections de trois pages au total : la première est consacrée à la réalisation de diagrammes de fréquence des valeurs de flux (la distribution est proposée sous une forme segmentée en classes) ; la seconde section consiste à choisir les flux à représenter, par exemple, les 2,5% de valeurs les plus élevées (De Jong et al., 2010 : 45-47) .

⁶ Le logiciel de cartographie de flux de W. Tobler (1979), intitulé Flowmapper© dont une présentation est encore disponible à l'adresse suivante : <http://www.csiss.org/clearinghouse/FlowMapper/> 9 dans sa version pour la gamme ArcGis n'est hélas plus maintenu.

⁷ La question de la réduction de la complexité de la carte statistique pour assurer sa lisibilité, et son efficacité, est un problème ancien et bien connu. L. Jéjou et J.-P. Deblonde (2012) ont d'ailleurs récemment revisité cette notion, dans le cadre de la cartographie chroplèthe.

1991 ; Zanin et Tremelo ; à celle de « *cartes de mouvement* » (« *movement mapping* ») (Bertin, 1967 ; Tobler, 1987) voire même de « *courants d'échanges* », de « *courants migratoires ou de migrations* » (Ravenstein, 1885 ; Courgeau, 1988).

Les concepts et notions de *flux*, de *mouvement*, de *courant*, de *d'échanges* sont-ils égaux ou synonymes ? Leur représentation cartographique est-elle similaire ? Si cela n'est pas le cas, pour quelles raisons ces termes semblent-ils substituables ? Quels sont les cas pour lesquels la carte décrit des flux et ceux pour lesquels elle décrit des échanges, des mouvements ?

Ces interrogations liminaires traduisent une difficulté globale d'appréhension de l'objet « carte de flux », justifiant la concomitance de dénominations présentant des significations variées (flux *versus* échange *versus* mouvement *versus* courant). Nous faisons l'hypothèse forte que **cette difficulté émane de l'éviction générale du tableau sous-jacent au processus de cartographie de flux. Autrement dit, que le type de matrice dont est issu la carte n'est jamais évoqué. Pourtant, c'est à partir du type de matrice que le type de carte (de flux, d'échanges, de mouvement, ...) pourra être déterminé et, par la suite, que le mode de représentation adéquat (du flux, de l'échange, du mouvement...) pourra être appliqué.**

Cette hypothèse justifie, à elle seule, le fait que l'objet de cette thèse soit la matrice de flux et non la carte de flux. En effet, nous apprendrons que la carte de flux correspond en fait à une représentation de la matrice parmi d'autres (chapitre 2) et que cette représentation correspond à un type de matrice en particulier (chapitre 3). Ne prendre en considération que la carte de flux serait réducteur, pour plusieurs raisons. D'abord parce qu'elle empêche de considérer d'autres modes de représentations qui lui sont éventuellement complémentaires. Ensuite parce que, comme nous l'avons déjà évoqué, les possibilités analytiques de la matrice sont variées. Ne pas les mentionner réduit ainsi l'éventail des possibilités cartographiques, pour les mêmes raisons que ne pas envisager les différents traitements applicables à un tableau d'information géographique réduit les possibilités de variation de la carte. Enfin, la matrice de flux étant formellement un tableau de contingence, lorsqu'elle est complète, elle peut être analysée – et donc cartographiée – selon différentes focales qui correspondent aux différentes parties du tableau : des *liens* et/ou des *lieux*.

C'est en réalité ma pratique professionnelle de la cartographie statistique qui m'a poussée vers ce questionnement sur les différents points de vue d'une matrice de flux. Cette expérience m'a conduite à soulever certaines difficultés d'ordre méthodologique et pratique qui concourent à la réalisation d'une carte de flux.

La première de ces difficultés est relative au caractère lacunaire des données portant sur les déplacements internationaux de populations. Nous n'allons pas approfondir ce point, car il s'agit d'un élément caractéristique pour ces flux mondiaux⁸ (Bahoken, 2007)⁹.

La seconde difficulté, davantage contextuelle, est liée à la compréhension de l'information. Les données de flux sont par essence complexes et nécessitent une spécialisation thématique pour pouvoir être analysées. Celles qui étaient alors disponibles dans un tableau faisaient état du résultat de la migration et non de la migration en elle-même. Cette précision est loin d'être anodine sur le plan cartographique. Elle signifie que ces données de flux, parce qu'elles sont situées sur une marge de la matrice, ne peuvent être représentées par des figurés linéaires – dont on verra qu'ils ne s'appliquent qu'au cœur de la matrice – quand bien même ces migrants se sont déplacés et que l'on connaît leur lieu d'origine (pays de naissance, par exemple) et leur lieu de destination (pays de résidence au

⁸ Fondées sur la déclaration des Etats, les données sur les flux internationaux de population ne sont pas, dans l'ensemble, exactes pour un pays donné, ni harmonisées, quand bien même elles sont collectées par un organisme unique, tel que par exemple, l'Office international des migrations (OMI).

⁹ L'ensemble de mes travaux intéressant l'analyse (cartographique) des flux est présenté à la suite de la bibliographie.

moment de la collecte des données). Ce constat prouve bien que l'objet « matrice de flux » est au centre du processus de cartographie de flux et non l'objet « carte de flux ».

Une troisième difficulté est relative à l'identification des références bibliographiques dédiées aux méthodes de cartographie de flux. Trois documents nous apparaissent révélateurs de la diversité de points de vue selon lesquels la problématique de la cartographie d'une matrice de flux peut être abordée (Cauvin, Reymond, 1984 ; Torricelli, 1997 ; Müller, 1991) :

- du point de vue du développement d'outils, afin de rendre opérationnelles les méthodes existantes ou d'actualiser d'anciennes procédures considérées comme désuètes (l'article de G.-P. Torricelli remet d'ailleurs au goût du jour une méthode de représentation développée au début des années 1970 par J. Bertin) ;
- du point de vue des méthodes de réalisation des cartes statistiques de flux, destinées en particulier aux atlas, dont les procédures (liées au traitement de données et à leur représentation cartographique) gagneraient à être formalisées, afin d'éviter erreurs de représentation et « *innocentes tricheries* », pour reprendre une expression de M. Monmonnier (1993), issue de son ouvrage *Comment faire mentir les cartes ?* ;
- du point de vue de l'exploitation des matrices de flux, mais à des seules fins de cartographie. L'exploitation des matrices de flux est, dans bien des cas, toujours envisagée dans le cadre de l'étape de modélisation spatiale des flux. Si celle-ci est nécessaire, elle n'est pas constitutive du processus de cartographie de flux. Nous montrerons que la modélisation des flux, comme d'autres méthodes de traitements de données, relève davantage de l'interprétation de la carte, de l'explication de ses motifs, que de la cartographie proprement dite. La thèse de M. Poulain (1981) *Contribution à l'analyse spatiale d'une matrice de migrations internes*, ne fait pas état des aspects liés à la représentation cartographique des résultats obtenus. L'auteur propose plutôt un examen minutieux des critères et paramètres des modèles appliqués aux matrices de flux de populations analysées. Ainsi, l'exploitation des matrices de flux dont il est question ici concerne l'utilisation de ses propriétés générales et de quelques nécessaires opérations que nous étudierons dans le contexte de l'analyse cartographique des flux – dès le chapitre 3.

Indépendamment de leur grand intérêt, ces trois textes issus d'une littérature académique, ne développent pas les aspects méthodologiques de la cartographie de flux, leur sémiologie. Comment expliquer cette situation apparemment secondaire de la carte de flux ?

Au sein de la littérature scientifique portant sur la cartographie statistique en général¹⁰ et/ou sur les traitements de données numériques¹¹, peu de références comportent une entrée spécifique aux flux (ou aux interactions spatiales). Lorsqu'elles présentent une section sur les flux, par exemple dans (Brunet, 1987 ; Dorling, 1991, 2012 ; Bertin, 1967 ; Zanin et Tremelo, 2003), celle-ci est généralement placée dans la partie de l'ouvrage consacré aux cartes qui sont difficiles à réaliser ou complexes. Ainsi, dans la section *Visages de la carte* de *La Carte mode d'emploi* de R. Brunet (1987), la carte de flux est associée à celle de réseau et placée à la dixième place parmi la vingtaine de propositions. Sa position dans le chapitre est également intéressante : elle est placée juste avant les cartes atypiques et complexes (carroyages, anamorphoses, ...). Dans la thèse de D. Dorling (1991), la cartographie des flux fait l'objet du chapitre sept intitulé *La Toile des flux (Cobwebs of Flows)* parmi les neuf qu'il propose, soit juste avant la présentation de modèles de cartes complexes réalisées en trois dimensions.

¹⁰ On peut notamment citer les références suivantes : André, 1980 ; Beguin, Pumain, 1994, 2003 ; Bertin, 1967, 1973, 1980 ; Bonin 1975 ; Bord, 1975 ; Brunet, 1987 ; Cauvin, Escobar, Serradj, 2008, tomes 1 et 3, Dorling, 1991 ; Le Fur, 2007 ; Le Fur & al., 2010 ; Rimbart, 1964, 1966 ; Joly, 1974, 1976 ; Monmonier, 1993 ; Palsky, 1996 ; Poitevin, 1999 ; Pumain, Saint-Julien, 1997, 2001 ; Weger, 1999 ; Zanin, Trémélo, 2003.

¹¹ Beguin, Pumain, 1994, 2003 ; Cauvin & al., 1987 ; Groupe Chadule, 1987 ; Cauvin, Escobar, Serradj, 2008 ; Pumain, Saint-Julien, 2001 ; Sanders, 1989 ; Vigneron, 1997.

J. Bertin (1967) présente *La représentation du mouvement dans le plan* dans le chapitre 8 de son ouvrage consacré à la *Sémiologie graphique*. Le manuel de A. Le Fur (2000) consacre la dernière page du dossier sur les modes de représentation cartographique, aux flux et aux réseaux (Le Fur, 2000 : 83). Enfin, le manuel de C. Zanin et M-L. Trémélo (2003), consacré à la conception cartographique des cartes univariées, présente dans sa fiche *Flèches et flux*, le mode de construction des flèches, mentionnant la spécificité de la carte qui « [...] implique de résoudre de nombreuses difficultés » (Zanin, Trémélo, 2003 : 173). Le document de C. Grasland et M. Baron (2006) utilisé pour l'Ecole d'été de Yaoundé, consacrée à l'analyse des flux et à la représentation cartographique, n'évoque pas ce dernier point : il présente la variation de la carte de flux en fonction de l'information résiduelle représentée. D'autres ouvrages sur les méthodes de la cartographie statistique de stocks, pourtant fort intéressants, ne présentent pas spécifiquement le cas de la représentation des flux ; c'est le cas, par exemple, de l'ouvrage de D. Poitevin (1999). Enfin, si le manuel de D. Pumain et T. Saint-Julien porte spécifiquement sur le traitement des interactions spatiales, il n'aborde que peu, voire pas, la question de leur représentation, c'est-à-dire la manière dont les cartes de flux présentées à titre d'illustration ont été réalisées. Serait-ce parce qu'il y aurait peu à dire sur la manière de réaliser des cartes de flux ? Pourtant, la cartographie de phénomènes d'accessibilité, qui peuvent être considérés comme une variante de la carte de flux (parce qu'ils sont focalisés sur une analyse des flux du point de vue du lieu de destination, dont on estime l'accessibilité depuis des lieux d'origine), font l'objet de plus de recherches (Cauvin, 1998 ; L'Hostis, 1997), mais ce n'est pas le sujet.

Aussi le travail que nous proposons sera t'il focalisé sur **les aspects liés à la cartographie du cœur de la matrice, celle qui porte sur la mise en relation des lieux que l'on représente traditionnellement par une flèche.**

Au-delà de cette déclaration d'intention, on peut s'interroger sur les raisons de cette position de la carte de flux dans les ouvrages qui en font état. Serait-elle liée au fait que la méthode soit banale, supposée connue même des non initiés ? Emettre cette hypothèse justifierait-il de porter les efforts sur les seuls aspects liés au traitement des données ? Les méthodes de représentation cartographique des matrices de flux sont-elles réellement maîtrisées de tous ? Si tel est le cas, comment expliquer le fait que leur réalisation apparaisse surtout dans certains travaux qui mettent en œuvre des modèles permettant d'expliquer ces flux, comme par exemple dans D. Pumain (1986), D. Pumain et T. Saint-Julien (1990) ou encore C. Dujardin (2001) ? En effet, étant donné que la cartographie de flux apparaît comme une composante de la modélisation des interactions spatiales, une étape de l'analyse, les solutions à apporter aux éventuels problèmes posés par sa construction seraient à rechercher prioritairement au moment du traitement des données numériques.

Quand bien même la cartographie statistique comporterait une part importante de traitement numérique, les solutions ne seraient-elles que d'ordre statistique ? Nous faisons l'hypothèse que **cette association modélisation-carte de flux n'est pas unique, qu'il est possible de modifier la figure sans toucher aux données numériques, c'est-à-dire en agissant sur les seuls aspects purement sémiologiques de la représentation, en réalisant une association sémiologie-carte de flux.** Si cette hypothèse est évidente, elle n'a pas fait l'objet, à notre connaissance, d'une validation empirique.

Cette thèse ambitionne ainsi d'apporter quelques clés de compréhension à la manière dont une carte de flux peut être construite, à partir de données disponibles dans une matrice et d'un fond de carte correspondant aux lieux d'origine et de destination, en amont de la modélisation qui n'est dnc pas l'objet de la thèse.

Nous proposons pour ce faire, une démarche d'ordre méthodologique.

Démarche mise en œuvre

La démarche proposée est de nature méthodologique, plutôt que technique ou technologique. En effet, nous avons observé que l'un des obstacles à la bonne appréhension des cartes de flux, à la difficulté de leur construction, est souvent attribué à celle du maniement d'un outil. De notre point de vue, cela relèverait plus d'une méconnaissance du corpus méthodologique pouvant être mis en œuvre que de l'utilisation d'un outil, comme l'avait d'ailleurs indiqué plus généralement P. Haggett à propos de la recherche en cartographie. « *La cartographie n'est pas seulement une représentation visualisée des localisations. C'est aussi l'expression synthétique d'une masse d'information dont le traitement et la traduction graphique exigent une recherche méthodologique d'ensemble* ». (Haggett, 1973 : 373).

Le coût général d'accès aux outils permettant une manipulation aisée des matrices de flux est, généralement, élevé pour qui n'y a pas été confronté pendant sa formation. Aussi, en l'absence de connaissances sur les fondements théoriques et méthodologiques de la procédure à appliquer, on est en mesure de se demander à quoi l'outil pourrait bien servir.

La difficulté de représenter des données matricielles conduit à une stratégie d'évitement de la carte de flux déjà mentionnée par D. Dorling (1991), au profit d'autres modes de représentation supposés plus faciles à mettre en œuvre, car les procédures sont implémentées dans les outils standards du marché. Cette stratégie n'est d'ailleurs pas propre à l'analyse géographique, elle a pu être identifiée dans d'autres domaines, par exemple en histoire par P. Rosental (1997), en informatique par Ghoniem et al. (2004, 2005) ou encore en biologie des populations par S. Lanco-Bertrand : « *Cet état de fait trouve son origine dans i) la difficulté en général d'obtenir des enregistrements suffisamment complets des organismes étudiés et ii) à la difficulté intrinsèque d'analyse d'un processus décrivant une occupation de l'espace-temps (Frontier et al., 2004) et variant donc dans au moins deux directions, l'une spatiale et l'autre temporelle [...] Turchin, 1998, va jusqu'à affirmer que ces difficultés pratiques et conceptuelles ont conduit à un comportement d'évitement de la part des écologistes* » (Lanco-Bertrand, 2005 : 127).

Le regain d'intérêt général observé pour les questions liées à la géovisualisation des graphes – dont on pourrait croire que la cartographie de flux pourrait tirer directement profit – est pourtant une illusion. En effet, les développements algorithmiques ayant conduit à des images spectaculaires, véhiculées par tous les media, sont réalisés dans le champ de l'informatique graphique (*InfoViz*). Si l'objectif est celui d'opérationnaliser le dessin de graphes, afin de révéler, selon E. Tufte (1990), une sélection de leurs caractéristiques structurelles qui seront formalisées graphiquement, ces images spectaculaires ne permettent pas, dans l'ensemble, de produire une connaissance thématique. C'est davantage la mise en scène de ces images qui explique leur popularité. L'outil permettant leur mise en œuvre présente un coût d'accès trop élevé et un aspect « boîte noire », opaque qui ne facilite ni leur appropriation ni leur démocratisation.

L'intérêt aujourd'hui renouvelé pour la visualisation graphique des données statistiques – des *data* – tient aussi et pour beaucoup à la technologie sous-jacente. D'après Fairbairn et al. (2001), cités par M.-J. Kraak, la visualisation des données dépasse, dans certains cas, le niveau de la création d'alternatives graphiques visuelles aux tableaux de valeurs. « *Elle concerne également les méthodes et techniques qui sont derrière le traitement des figures, l'environnement [informatique] dans lequel les données sont utilisées (i.e. l'interface) et la réponse à la question 'Cela marche ?'* » (Kraak, 2002 : 393), ce qui constitue un réel investissement. Le résultat revêt alors un caractère précieux qui tient davantage aux prouesses techniques et technologiques de quelques-uns, que de la capacité analytique de la figure qui n'est pas toujours assurée. Ce constat trivial tient surtout à une déconnexion récurrente – déjà mentionnée par de nombreux auteurs – entre les développeurs d'outils ou d'applications et les thématiciens ou les cartographes se positionnant de plus en plus à l'interface de ces deux profils.

Les méthodes proposées pour cartographier des flux ne sont pas spécifiques à un outil particulier, mais le précèdent. L'outil correspond au moyen permettant la mise en œuvre de la méthode et il est généralement interchangeable, ce qui n'est pas le cas des méthodes. En informatique, l'algorithmie permet de formaliser des méthodes et de développer des outils. Cela étant, on ne peut pas dire que les outils soient tous égaux ou que la méthode ait toujours vocation à être opérationnalisée dans un outil. La démarche proposée dans cette thèse souhaite s'inscrire dans le cadre d'une recherche reproductible, aussi considère-t-elle comme secondaires les questions liées aux outils. Ces derniers ne seront donc pas examinés en profondeur. Toutefois, notre démarche se voulant *générale* et *généralisable* et applicable à tous les types de matrices de flux, nous proposerons néanmoins quelques programmes développés sur la plateforme RStudio©.

Nous savons aujourd'hui que la cartographie des matrices de flux souffre d'un certain nombre de contraintes et de difficultés essentiellement d'ordre méthodologique. Trouver la solution à l'une d'entre elles consiste à appliquer plusieurs façons de procéder, aux différentes étapes de la conception cartographique : lors d'un¹² traitement des données, du tracé des figurés, du choix d'une variable visuelle, de teintes, ...

Cette *Contribution à la cartographie d'une matrice de flux* est ainsi composée de deux grandes parties : la première dresse l'état de la question des contraintes et des difficultés méthodologiques propres à la cartographie de flux, la seconde propose une sélection de solutions à certains des problèmes identifiés.

Le caractère général de cette contribution consiste à laisser l'aspect thématique des données le plus indéterminé possible afin de n'identifier (en première partie) ou de ne rechercher (en seconde partie) que les solutions qui s'appliquent à la cartographie des données de flux. Cependant, pour expliciter au mieux l'application des différentes solutions, nous aurons recours à des exemples variés illustrant les procédures ainsi qu'à des données réelles permettant de les valider et d'assurer leur caractère généralisable.

Les méthodes proposées sont par nature non exhaustives et incomplètes ; elles seront proposées pour résoudre tout ou partie d'un problème précis. Nous supposons ainsi connus les procédés de cartographie de données de stocks ainsi que les méthodes qui concernent la préparation du fond de carte. Cela nous permet de nous concentrer sur les données de flux. L'investigation menée dans la première partie permet de connaître les fondements des méthodes de cartographie existantes, ainsi que leurs contraintes et les éventuelles difficultés de leur mise en œuvre. La réaliser nous a surtout donné la possibilité d'analyser notre propre pratique, de formaliser notre démarche de compréhension des données matricielles pour mieux concevoir la carte de flux.

Cette connaissance correspond à notre représentation des questions posées, elle émane de la formalisation de notre pratique dont la compréhension s'inscrit indubitablement dans un corpus disciplinaire : celui de la géographie humaine (des sciences du territoire) étendu aux disciplines connexes du champs des sciences humaines et sociales qui mobilisent des matrices de flux à des fins de cartographie (démographie, sociologie, statistiques) ou comme objet d'étude privilégié (biologie des populations, par exemple).

¹² Ces programmes sont inspirés des travaux réalisés dans le cadre des groupes de travail flux, matrices, réseaux et <http://groupefmr.hypotheses.org> et ElementR <http://elementR.hypotheses.org> [liens vérifiés le 06/08/2015].

Cadrage théorique du sujet

Notre démarche s'inscrit dans la mouvance théorique et quantitative de la géographie humaine qui considère exacte une boutade de P. Pinchemel selon laquelle : « *A la limite, tout n'est que flux en géographie* »¹³.

Sur un registre plus classique, nous avons adopté la grille d'analyse proposée par C. Grasland (2009) et l'avons adaptée au cas de la cartographie des interactions spatiales, des mobilités individuelles ou agrégées. De cette manière, notre objectif de cartographie d'une matrice de flux ne saura être dissocié de celle de ses lieux, dans le cadre d'une approche que nous qualifierons de *liens-lieux*. Cela nous conduit à ancrer notre analyse dans le courant de la géographie qui considère le rôle général joué par l'espace dans l'expression des flux de différents types qui s'expriment entre des lieux, nécessairement distincts et distants ; l'ensemble de ces lieux formant dans la réalité un territoire en *mouvements*. L'espace géographique n'étant ni abstrait, ni donné en l'état et par conséquent ni dénué de sens (il est façonné par les sociétés qui le pratiquent et l'investissent) : « *En définissant l'espace géographique comme un ensemble de lieux et de relations qui s'établissent entre ces lieux, on place au centre la notion d'interaction spatiale [...]. Mouvements, échanges, ou influences, entre les lieux résultent des identités et des rôles des acteurs qui, occupant une position dans l'espace géographique, vivent sous la contrainte expresse de l'échelle avec d'autres acteurs localisés ailleurs [...]. L'ensemble de ces déplacements, massifs et en croissance continue, donne forme et mouvement à l'espace géographique* » (Saint-Julien, 2010 : 16).

Cet ancrage géographique de notre réflexion cartographique conduit à inscrire nos travaux à la marge de deux axes : le premier est celui du traitement des données matricielles et le second, celui de leur représentation graphique. Si l'étape de traitement des données de flux est issue de travaux réalisés par des démographes tels que D. Courgeau (1980), elle a aussi fait l'objet d'une appropriation par plusieurs géographes parmi lesquels W. Tobler qui s'intéresse à la cartographie des matrices de flux vers la fin des années 1970. L'étape de représentation des données statistiques qui portent sur des flux a été initiée dès la fin du XIX^{ième} siècle par des ingénieurs européens, parmi lesquels H. D. Harness (1836), C.-J. Minard (1869) et M. Sankey (1898). Elle n'a toutefois été formalisée qu'à la fin des années 1960 par J. Bertin (1967) qui développe une véritable recherche théorique en cartographie, sur les procédés graphiques et les aspects cognitifs de leur compréhension, alors que cette dernière était jusque là cantonnée aux aspects pratiques de la réalisation de cartes.

C'est pourquoi, nous émettons l'hypothèse que **les solutions (apportées) aux difficultés posées par la cartographie des matrices de flux sont d'abord à rechercher dans les travaux de J. Bertin et W. Tobler.**

J. Bertin formalisa en 1967 l'ensemble des recherches qu'il avait engagé depuis les années 1950, dans un ouvrage remarquable de sémiologie graphique, qui sera plusieurs fois édité et complété. Presque immédiatement après sa parution, W. Tobler en proposa une analyse critique (1968). Tout en reconnaissant les apports indéniables de l'auteur, en termes de sémiologie graphique, au point de considérer son œuvre comme une référence, il regrettait l'absence de références bibliographiques, le nombre limité des traitements aux quatre opérations mathématiques de base et la brièveté des démonstrations. L'absence de références est d'ailleurs un des principaux reproches fait à J. Bertin qui élude, consciemment ou inconsciemment, les travaux antérieurs. Par exemple, les procédés graphiques de la représentation des *mouvements* que l'auteur évoque¹⁴ ont, en effet, été posés pour l'essentiel dès le XIX^{ième} siècle (Harness, 1837 ; Ravenstein, 1885,1889 ; Minard, 1844-1869). J. Palsky et M.-C.

¹³ Cette jolie phrase a été prononcée lors d'un séminaire du Diplôme d'études approfondies : Analyse théorique et épistémologique en Géographie (DEA ATEG) en 1987, selon le témoignage de C. Grasland, alors étudiant.

¹⁴ Nous y reviendrons dans le chapitre 4.

Robic ont, par ailleurs, mis en exergue la sémiologie d'avant J. Bertin (Palsky et Robic, 2000). Il faut toutefois préciser que nombre de notes et de manuscrits de J. Bertin ne sont pas publiés – ils sont accessibles à la Bibliothèque Nationale de France¹⁵ depuis son décès en 2010. A *contrario*, les travaux de W. Tobler, articles, communications orales et notes de recherches non publiées ont fait l'objet d'une mise en ligne systématique et de compilations (Cauvin, Grasland, Schlumberger, 2000).

De fait, si l'on s'en tient à ses deux ouvrages principaux, les travaux de J. Bertin intéressent majoritairement les aspects liés à la représentation graphique des données spatialisées ainsi qu'aux aspects liés à leur perception. Cependant, les résultats de certains traitements qu'il propose sont identiques à ceux qui auraient été obtenus à l'aide de modèles mathématiques (Bertin, 1980).

Les travaux de W. Tobler, auxquels nous faisons référence ici, portent sur la modélisation et la représentation des interactions spatiales, essentiellement sous la forme de carte. Il s'agit de résultats qui intéressent les recherches de l'auteur menés dans ses deux premières périodes : « [...] –*La première période correspond à la mise au point d'une cartographie assez développée, associée à un modèle permettant d'estimer le pouvoir attractif des lieux – une deuxième étape aboutit à la production d'un modèle simple et opérationnel utilisant le calcul matriciel, partant de données plus réduites ce modèle permet de retrouver les mouvements entre les lieux et l'attractivité de ces lieux* » (Cauvin, Reymond, 1991: 468). Il est intéressant de noter que le premier¹⁶ problème que W. Tobler cherche à résoudre est d'ordre cartographique. Il concerne la représentation de mouvements, compte tenu de la position de couples de lieux et de la valeur des flux qui les caractérisent. Dans la seconde période, W. Tobler utilise la cartographie pour illustrer les résultats obtenus dans le champ de la modélisation¹⁷ des interactions spatiales. Cela conduit l'auteur à transférer les méthodes de l'analyse matricielle au cas des matrices de flux, notamment à proposer une méthode de décomposition de l'information de flux qu'elle contient¹⁸.

Si J. Bertin ne référence pas particulièrement ses travaux, W. Tobler emprunte les siennes à de nombreuses disciplines, contribuant ainsi à renouveler profondément les aspects conceptuels et méthodologiques d'une analyse spatiale du *mouvement*. L'un des quatre volumes¹⁹ reprenant la majorité de son travail, composé de 34 références, intéresse directement la cartographie des mouvements (Tobler, 2000).

J. Bertin et W. Tobler ont ainsi en commun une réflexion d'ordre théorique et méthodologique sur la représentation cartographique de différents flux. Si les travaux du premier sont davantage focalisés sur

¹⁵ La liste des documents consultables de J. Bertin est disponible en ligne, d'une part, sur le site *data* de la BNF, aux adresses suivantes : http://data.bnf.fr/11985023/jacques_bertin/ et <http://data.bnf.fr/documents-by-rdt/11985023/70/page1> en particulier dans les pages de l'atelier expérimental dédiées à l'auteur et disponibles sur ce même site. D'autre part, sur le site de l'Atelier de cartographie de Sciences Po réalisé avec la collaboration de l'auteur, à l'adresse suivante : http://cartographie.sciences-po.fr/en/la_graphique_jacques_bertin2 [liens vérifiés le 02/08/2015].

¹⁶ Le second problème que W. Tobler cherche à résoudre, dans les recherches de sa première période, concerne l'explication de ce qu'il qualifie d'attractivité des lieux (Cauvin, Reymond, 1993 : 468). Ce second problème n'est pas cartographique, il intéresse plutôt les aspects mathématiques de l'analyse des flux qui sont au cœur de l'investigation de l'auteur, elle précède l'étape de cartographie des attractivités qui intervient dans un second temps.

¹⁷ Ses recherches portent sur la mise au point de la version additive du modèle gravitaire, parallèlement à son inversion qui consiste à estimer les valeurs de flux à partir des distances et des sommes marginales de la matrice.

¹⁸ Nous repartons de cette proposition dans le Chapitre 3 pour proposer une formalisation de l'approche classique de cartographie de flux.

¹⁹ C. Cauvin et C. Grasland, avec la collaboration de R. Schlumberger, ont réuni au sein de quatre volumes, différents textes publiés, non publiés et notes de recherches que W. Tobler réalisa entre 1959 et 2000. Les travaux réalisés après 2000 sont disponibles en ligne, sur la page personnelle de W. Tobler : <http://www.geog.ucsb.edu/~tobler/> [lien vérifié le 02/08/2015].

les aspects graphiques et cognitifs, ceux du second approfondissent les aspects mathématiques qui concourent également à faire varier la figure.

D'autres auteurs ont, bien entendu, influencé notre réflexion, mais certainement dans des proportions moindres. Nous situons ainsi notre raisonnement dans une zone d'interface entre la sémiologie cartographique, la géovisualisation des données relationnelles et l'analyse spatialisée des données, au sein d'une démarche d'ensemble valorisant la complémentarité des approches statistiques et (carto)graphiques, la fertilisation croisée de W. Bunge (1962). D'après J. Tukey (1977), dont l'ouvrage consacré à l'exploration et la visualisation interactive de données statistiques est un bon exemple de cette complémentarité²⁰ entre aspects statistiques et aspects liés à la représentation (carto)graphique de données numériques, est un puissant outil d'analyse selon différentes perspectives. Complémentaire aux traitements statistiques, elle permet d'explorer les données, de susciter des questionnements, de valider des hypothèses visuellement pour aboutir à une connaissance. Elle contribue ainsi à éclairer les données et ce d'autant plus que la figure est efficace.

Plan général

Notre cadre méthodologique ayant aussi une visée pédagogique, nous avons couplé les deux grandes parties déjà présentées à une dialectique : identification des problèmes (chapitres 1 à 4) | propositions de solutions théoriques (chapitre 6) et méthodologiques (chapitres 5, 7, 8 et 9).

La première partie vise à se familiariser avec la construction des données matricielles (chapitre 1), les différentes manières de les représenter graphiquement (chapitre 2) en particulier sous une forme cartographique. Elle conduit à une formalisation des pratiques actuelles qui prend la forme d'une typologie des méthodes de cartographie de flux, en fonction du type de matrice (chapitre 3). Elle permet aussi une mise en perspective des différentes possibilités de cartographie existantes depuis 1836, tant sur le plan méthodologique que sémiologique (chapitre 4) ; l'identification de défis qu'il convient de relever, présentés sous la forme de problèmes d'ordre méthodologique et de contraintes graphiques.

La seconde partie propose un éventail de solutions non exhaustives et opérationnelles. Elles permettent de résoudre une sélection de problèmes liés à la complexité visuelle : graphique et analytique de la carte de flux. Les solutions sont d'abord d'ordre sémiologique, sans action sur les données (chapitre 5) et conduisent à renouveler l'apparence des flux. La discussion sur la signification de la carte de flux, en particulier au regard de celle du mouvement (chapitre 6), nous conduit à intégrer un critère spatial (chapitre 7) dans le raisonnement cartographique, puis à envisager différents cas de sélection des flux à cartographier, en mobilisant un critère global unique (chapitre 8) ou un critère local, focalisé sur les lieux afin de les hiérarchiser selon leur attractivité et de régionaliser l'espace relationnel issu des flux dominants (chapitre 9).

²⁰ De courtes vidéos des travaux de J. Tukey (1915-2000) sont disponibles à l'adresse suivante : <http://stat-graphics.org/movies/prim9.html> [lien vérifié le 02/08/2015].

PREMIERE PARTIE

Etat de l'art des méthodes de cartographie de flux. Enoncé des problèmes, des contraintes graphiques et des difficultés d'ordre méthodologique

INTRODUCTION DE LA PREMIERE PARTIE

« Ce dernier devra prêter d'abord attention à leur principe de construction, faute de quoi, il pourrait aussi bien regarder une feuille blanche ».

William Playfair (1805).

Cette première partie a pour objectif d'énoncer les problèmes, les contraintes graphiques et les difficultés d'ordre méthodologique liées à la représentation cartographique d'une matrice de flux. Nous proposons pour cela un état de la question qui articule des éléments d'ordre structurel, liés aux données statistiques dont on dispose pour l'analyse, aux aspects méthodologiques de la conception d'une carte de flux. Cet état de la question nous conduira à identifier les différents procédés de représentation de flux, il nous conduira à l'identification -au cours des quatre chapitres qui composent cette première partie - d'une série de contraintes et de difficultés d'ordre méthodologique que nous allons récapituler en conclusion.

Le premier chapitre présente l'objet d'étude : la matrice de flux. Nous examinerons les valeurs disponibles dans la matrice ainsi que ses propriétés générales. L'information matricielle de flux agrégés de lieux à lieux est illustrée par la comparaison de matrices de flux migratoires et commerciaux. Leur élaboration étant le résultat d'une série d'opérations statistiques d'anonymisation, de filtrages et d'agrégations spatio-temporelles de déplacements individuels, nous démontrerons l'incomplétude qui en résulte, à partir d'un exemple fictif mais plausible. L'exploitation du récit biographique d'une femme prénommée Marie (1940-1960), nous permettra de reconstituer une (voire plusieurs) matrice(s) de ses mobilités résidentielles, tout en présentant les principaux concepts et l'essentiel des notions impliqués dans la construction des valeurs de flux. Cet exemple permettra de présenter les concepts et notions impliquées dans la formation des valeurs de flux (Fij), ainsi que les contraintes liées au dispositif utilisé pour la mesure. Le problème de l'incomplétude des flux identifiés, la question de leur représentation graphique pourra être envisagée.

Le second chapitre représente la matrice selon trois modèles graphiques : la matrice ordonnée, le graphe et la carte de flux. L'objet carte de flux qui correspond formellement à un graphe valué spatialisé est élevé au rang de modèle, en raison du rôle joué par la composante spatiale, tant dans la localisation des lieux que dans les interactions qui les caractérisent. Nous montrerons ainsi que ces trois modèles conduisent à une transformation en graphique des données numériques. Dans le premier cas, la transformation tient un recodage ordonné des valeurs en niveaux colorés, sans affecter la structure d'ensemble de la matrice. La transformation est double dans le cas du graphe : elle porte sur les valeurs mues en objet graphique (des liens et des nœuds) qui sont placées dans un plan quelconque, à la suite d'un repositionnement optimisé des points. La carte de flux projette, quant à elle, ces mêmes objets graphiques (des liens et des lieux) dans un espace assorti d'une métrique (généralement

géographique). Ces trois modèles sont illustrées par une matrice de migrations internes, celles des flux interdistricts de l'ex-Tchécoslovaquie (Grasland, 1994 : 439). La difficulté de représenter et de percevoir le résultat obtenu par la cartographie directe d'une matrice de flux, nous conduira à définir cette complexité graphique. Nous montrerons ainsi que *l'effet spaghetti* est issu de la conjonction d'un effet de *surcharge graphique* (densité de figurés) et d'un effet de *couverture spatiale* (densité de teinte, liée aux dimensions géométriques).

Le troisième chapitre examine le processus de cartographie directe d'une matrice de flux, en focalisant l'attention sur le cœur de la matrice (Fij). Nous présenterons dans un premier temps les principales approches statistiques et matricielles, liées au traitement des données de flux, formant le corpus actuel de la cartographie de flux. Nous adopterons ensuite la posture de W. Tobler (1979 1982), dont le cadre général d'analyse nous permettra de présenter une typologie des méthodes de cartographie de flux. Celle-ci résulte d'une procédure de décomposition de la matrice de données observées (Fij) en deux composantes : (Fij⁺) et (Fij⁻), qui se distinguent en fonction de leur caractère symétrique ou non au regard à la diagonale principale. Nous montrerons ensuite qu'à ces différentes matrices correspond en réalité une et une seule méthode de représentation cartographique, selon que le figuré linéaire pour représenter ces liens, (Bertin, 1967 ; Zanin, Tremelo, 2003) est orienté ou non. La typologie proposée, selon une cartographie conjointe des lieux et des liens comme suggéré par (Grasland, 2009), propose un cadre méthodologique général, fonctionne sur l'ensemble des matrices de flux. Ayant été formalisé au regard de notre pratique, il sera nécessaire de le valider en examinant les différentes figures de la carte de flux disponibles dans la littérature.

Le quatrième chapitre prend la forme d'une rétrospective où l'on présente l'évolution au cours du temps de l'image cartographique des flux. L'évolution des méthodes et procédés cartographiques étant le reflet d'une époque, les innovations théoriques et conceptuelles sont regroupées en six périodes de 1836 à 2013. Elles sont structurées autour de l'apport du ou d'un ensemble d'auteurs, géographes-cartographes, figures emblématiques et génies de la cartographie des flux et des réseaux spatiaux. A l'échelle locale, c'est le cas par exemple de E.G. Ravenstein qui signe en 1885 la première carte de flux réalisée avec des flèches et, bien sûr, de C.-J. Minard (1844-1877) et quelques autres. A l'échelle mondiale, se distingue notamment le génie de E. Raisz (1944), non seulement pour sa vision globale et polaire des flux et des réseaux, mais aussi parce qu'il publia en anglais le premier manuel portant sur des méthodes de cartographie. Cette présentation chronologique des innovations cartographiques atteint son apogée avec les travaux de W. Tobler (1981), qui conduisent notamment à une formalisation de l'approche empirique de E.G. Ravenstein, avant une nouvelle ère marquée par la carte de P. Butler (2010), décrivant sur un planisphère les liaisons mondiales des abonnés à un célèbre réseau social, caractéristique de l'ère contemporaine de l'*InfoViz*. Les différentes innovations sont toujours appréhendées au regard de la typologie des méthodes présentée dans le chapitre 3, ce qui permet de valider son caractère opérationnel, avant que ne soient récapitulées dans un second temps, les principes de sémiologie de la carte de flux. Ces derniers conduisent à une identification minutieuse d'une multitude de contraintes graphiques liées au tracé des différentes composantes du figuré linéaire.

A l'issue de ces quatre chapitres, nous identifierons quatre familles de problèmes ou de contraintes, résolus ou non, qui traduisent la difficulté de cartographier des flux.

1^{er} CHAPITRE. La construction de la matrice de flux : incertitudes de mesure et problèmes d'agrégation

Des données de flux fournies par « [...] 'paquets' [...] *incommodants et anachroniques qui posent des problèmes extrêmement ardu de cartographie et d'interprétation* ».

Peter Haggett (1973).

Introduction du chapitre

La réalisation d'une carte de flux résulte de l'exploitation des données disponibles dans un tableau particulier, la matrice de flux. Celle-ci contient des données statistiques, les valeurs de flux observés, et des informations d'ordre géographique qui portent sur une zone d'étude formée de lieux d'origine et de destination des flux. L'information géographique fait en général référence à un maillage territorial. Ce maillage forme une partition incomplète de l'espace puisque sur une carte de flux, on exclut souvent les espaces non urbains de l'analyse. L'espace est ainsi défini par N objets géographiques (i), éventuellement caractérisés par des attributs sémantiques (z). Il s'agit, par exemple, des communes d'un département et de leur effectif de population. L'information géographique porte également sur les couples (i,j) de lieux d'origine (i) et de destination (j) des flux, également caractérisés par des attributs (z) qui leur sont propres tels que la distance kilométrique qui les sépare, par exemple. Ces données géographiques levées sur un espace d'étude sont parfois rendues disponibles sous une forme vectorielle. Il s'agit du fond de carte qui, lorsqu'il est associé à une table de données individus*variables stocke également les caractéristiques des lieux et devient une matrice d'information géographique. L'information relevée sur le terrain, lorsqu'elle porte sur des flux, est archivée par ailleurs dans une *matrice de flux*²¹. Celle-ci décrit la valeur d'une relation intrinsèque aux couples de lieux concernés, en l'occurrence ici, un flux.

Dans notre cadre d'analyse des mobilités spatialisées, la matrice de flux présente différentes valeurs numériques, généralement quantitatives discrètes. Ces valeurs représentent le flux défini au sens général comme la quantité qui s'est déplacée depuis un lieu d'origine vers un lieu de destination, pendant une période donnée. Si les données de la matrice qui portent sur les caractéristiques spatiales et sémantiques des couples de lieux concernés sont aisément identifiables, il n'en est pas de même pour celles qui portent sur la quantité qui s'est déplacée. On peut même considérer qu'elle présente un aspect « boîte noire » qui nécessite d'être dévoilé, au risque d'erreurs d'interprétations des résultats qui peuvent s'avérer préjudiciables pour l'analyse et l'interprétation des flux.

²¹ Les termes ou groupes de termes placés en italique sans guillemets correspondent à des notions et des concepts fondamentaux pour la thèse. Leur définition est présentée dans le Glossaire, situé en fin de document.

Cette incertitude sur la valeur des flux disponibles dans une matrice tient à la nature multidimensionnelle de l'information relationnelle. La mesure des flux implique, en effet, le recours à plusieurs référentiels articulés : un référentiel spatial, le maillage de la zone d'étude ; une composante temporelle, la période d'observation du flux ; une composante thématique qui concerne le type de flux dont il est question. Ces trois référentiels comportent, par ailleurs, des acceptions qui leur sont propres. La composante temporelle, par exemple, peut être appréhendée comme une période $[t_1 - t_2]$ ou bien comme un instant (t).

« *b- On doit donc, dès que l'on veut formuler une généralité sociologique, revenir sur les conditions de production de l'information qu'on utilise pour les incorporer à l'énonciation des constats historiques, puisque la connaissance de ces conditions est indissociable du sens de l'énonciation de l'objet²² qu'elles ont permis de construire. C'est là, si l'on entend garder sa pertinence empirique à l'énonciation, le seul moyen d'énoncer généralement sur des séries d'objets construits différemment.* »²³ (Passeron, 1991 : 114).

Ce premier chapitre a pour objet de présenter la matrice de flux (Fij) en tant qu'objet d'étude devant faire l'objet d'une cartographie, de découvrir le principe de sa construction, d'explicitier en particulier ce à quoi correspond la valeur (Fij) portée au croisement des (i) lignes et des (j) colonnes qu'elle décrit. Cette valeur n'a, en effet, aucune signification directe si elle n'est pas appréhendée au regard de la procédure qui a permis de la générer, depuis la phase de collecte des données jusqu'à leur archivage dans la matrice. La présentation de la structure lignes*colonnes d'une matrice de flux quelconque nous conduit, dans une première section, à nous interroger sur le sens à donner aux valeurs qu'elle représente. Elle est illustrée par deux exemples, celui d'une matrice de flux migratoires et celui d'une matrice de flux commerciaux dont l'apparence et la ressemblance des valeurs qu'elles présentent sont tout aussi fortes que leurs différences sont importantes et, par conséquent, fondamentales. Cette comparaison justifie l'investigation sur la manière dont les flux sont mesurés. Elle nous conduit à présenter, dans une seconde section, la manière dont les valeurs de flux sont mesurées ainsi que différentes méthodes de collecte des flux migratoires et commerciaux archivés dans une matrice de flux. La donnée transmise par les organismes de collecte n'étant que, dans de très rares cas, celle qui a été mesurée initialement, nous présentons dans une troisième section, les procédures qui conduisent à la génération de valeurs de flux agrégés à partir d'une collecte réalisée au niveau élémentaire des déplacements. Pour ce faire, nous illustrons le cas général à partir de l'exemple de flux migratoires que nous enrichissons grâce au recours à une approche graphique des différents concepts mis en œuvre.

1.1 Présentation de la matrice de flux

Au préalable de leur représentation graphique, les données qui portent sur des relations entre des entités sont archivées dans un tableau croisé nommé matrice (Fij). Celle-ci est composée de m lignes et de n colonnes qui correspondent respectivement aux entités d'origine (i) et de destination (j). Le premier terme est toujours associé par convention aux lignes (\rightarrow) et le second aux colonnes (\downarrow). La structure de la matrice décrit ainsi une grille formée de (i) lignes, de (j) colonnes et de N ($N=m*n$) cellules portées en leur croisement qui décrit une structure particulière des relations observées.

²² Terme porté en italique dans le texte d'origine.

²³ Ce paragraphe est précédé d'un autre qui est le suivant : « *a- On ne peut, dans les sciences sociales, rester dans le cadre d'un raisonnement purement expérimental strict qu'en oubliant momentanément les conditions du prélèvement d'information qui permettent cet exercice : désignation du contexte dans lequel ont été mesurées les interactions entre les variables traitées et que commande le droit à la généralisation ou comparaison ; énumération des conditions méthodologiques de la construction de l'information qui subordonnent le sens de l'information aux formes techniques ou linguistiques de son recueil : succession des choix d'enquête, d'échantillonnage, de nomenclature et de traitement* » (Passeron, 1991 : 114).

1.1.1 Structure et propriété générale de la matrice de flux

La structure générale de la matrice prend la forme d'un tableau présenté indifféremment sous la forme d'une matrice individu*individu ($m*n$) ou bien d'une liste de N valeurs qui comporte au moins trois colonnes (3c), comme illustré par la Figure 1-1.

Figure 1-1. La structure générale d'une matrice

Une matrice ($n*m$)					Une liste (3c)		
$i \backslash j$	A	B	C	D	i	j	Lij
A	0	1	1	1	A	A	0
B	1	0	1	1	A	B	1
C	1	1	0	1	A	C	1
D	1	1	1	0	A	D	1
					B	A	1
					B	B	0
					B	C	1
					B	D	1
					C	A	1

avec :

- A : identifiant du lieu
- 1,0 : valeur d'adjacence (Lij)
- 1 (existence d'un lien)
- 0 (absence de lien)

La liste est une forme particulière de tableau (individu*variable) puisque deux colonnes sont réservées aux couples d'individus ou d'entités mis en relation. C'est pourquoi elle est formée d'au minimum trois colonnes qui correspondent respectivement aux identifiants des lieux d'origine (i), de ceux destination (j) et à l'existence d'un lien (Lij). Pour des facilités de lecture, il est recommandé de procéder à un tri croissant de la liste, selon les (i) lignes puis les (j) colonnes tel que présenté sur la partie droite de la Figure 1-1. Les deux formes de tableau de la Figure 1-1 sont strictement équivalentes et le passage de l'une à l'autre ne pose pas de problèmes particuliers. Il peut être réalisé à l'aide des tableurs²⁴ actuels ; de convertisseurs²⁵ indépendants ou à l'aide de quelques lignes de commandes²⁶ sous RStudio®, voir Annexe R-1.

Le croisement d'une ligne et d'une colonne (i, j) correspond à une cellule ou case de la matrice, généralement assortie d'une valeur. La valeur est donc placée à la i -ième ligne et à la j -ième colonne. Cette cellule notée (F_{ij}) renseigne, soit sur l'existence d'une relation entre les $i*j$ entités où $i=1, \dots, m$ et $j=1, \dots, n$, soit sur la valeur de cette relation. Lorsque les valeurs sont binaires (0,1), la matrice exprime l'existence d'une relation (Lij). Elle est alors qualifiée de *matrice d'adjacence* – voire de *connexité* – où la modalité '1' indique l'existence d'un lien et la modalité '0' traduit son absence. La matrice présentée sur la Figure 1-1 est une matrice d'adjacence.

²⁴ Les outils de type « tableau croisé dynamique » disponibles pour Excel® permettent de passer du format de liste au format de matrice, mais non l'inverse.

²⁵ On peut mentionner l'outil « Convertisseur de matrices », développé par I. Brémond (CNRS, UMR Unité Mixte de Recherche (UMR) 6590 Espaces et Sociétés (ESO), ne fonctionnant que pour les matrices carrées ; aussi le convertisseur développé par les ingénieurs de l'Unité Mixte de Services (UMS) 2414 Réseau Interdisciplinaire pour l'Aménagement du Territoire Européen (RIATE), sous RShiny, à l'adresse suivante : <http://riate.shinyapps.io/matrices/> [Lien vérifié le 02/08/2015].

²⁶ Sous RStudio®, pour transformer une matrice en liste, il convient d'utiliser la fonction « melt » et inversement, les fonctions « acast » ou « dcast ». Un exemple de leur mise en œuvre est notamment proposé dans l'Annexe R-1 dédiée à la transformation des matrices fermées non carrées en matrices carrées (où le nombre de lignes est égal au nombre de colonnes).

Dans certains cas, cette matrice d'adjacence est valuée, ce qui signifie que la cellule est assortie d'une valeur quantitative mesurable et définie positive. Cette valeur est soit discrète, si elle ne prend qu'un nombre fini de valeurs, soit continue, si elle en accepte une infinité dans \mathbb{R} . Ce sont ces matrices d'adjacence valuées que nous appelons matrices de flux. La Figure 1-2 présente la matrice de flux fictive que nous avons élaborée à des fins d'illustration des différentes notions et concepts, des traitements de données et de leurs résultats cartographiques. Cette matrice de flux sera, de fait, régulièrement mobilisée, tant dans la première partie de la thèse que dans la seconde partie, pour illustrer les différentes familles de solutions que nous proposons.

Figure 1-2. Une matrice de flux

		j			
		A	B	C	D
Fij	A	-	8	5	1
	B	6	-	3	2
	C	10	4	-	1
	D	1	2	3	-

avec :

- i : lieux d'origine
- j : lieux de destination
- A : identifiants des lieux
- 8 : valeur du flux (Fij) de (A,B)

Une matrice d'adjacence ordinaire peut ainsi être considérée comme un cas particulier de matrice de flux qui a fait l'objet d'une transformation, de type « binarisation », afin de ne renseigner que sur l'existence (ou non) d'une relation entre deux entités. L'information d'une matrice d'adjacence est, de fait, plus pauvre²⁷ que celle d'une matrice de flux, si l'on souhaite l'appréhender dans le champ de l'analyse des interactions spatiales, car elles ne contiennent qu'une information sur l'existence d'un lien entre des lieux, elles ne décrivent pas les quantités qui se sont déplacées. Ainsi, dans notre cadre d'analyse des flux, la *matrice d'adjacence* (L_{ij}) correspond à une dégradation de l'information de flux (F_{ij}) de départ. C'est pourquoi il ne sera plus question que de matrice de flux, sauf dans les cas particuliers qui seront signalés.

Nous restreignons toutefois notre propos au cas des matrices de flux où les lieux d'origine (i) sont identiques (j) aux lieux de destination, quelle que soit leur appellation. Les matrices concernées sont des matrices fermées, parfois appelées *matrices carrées*. Quoiqu'il en soit, ce choix nous permet de ne pas limiter d'emblée notre champ d'analyse à une sélection de lieux, en considérant tous ceux qui définissent l'espace d'étude. Nous conservons ainsi le cadre théorique général des N^2 possibilités de relations entre les N entités de la zone d'étude.

Le nombre N définit les dimensions de la matrice, elle est dite d'ordre ($i*j$) ou de rang N, correspondant au nombre d'entités considérées, c'est-à-dire au nombre de lieux de la zone d'étude. Dans le cas d'une matrice *carrée*, N correspond au nombre de lignes et de colonnes. La matrice de flux de la Figure 1-2 est ainsi carrée, d'ordre 4.

La matrice de flux est, en pratique, assimilée à un tableau croisé particulier : le tableau de contingence (Chadule, 1987 : 77), ce qui lui confère des propriétés particulières. La propriété d'additivité en ligne et en colonne permet, par exemple, de générer des indices, tels que le total des flux émis par un lieu (i) précis, égal à la somme en ligne notée (O_i) ; ou le total des flux reçus par un lieu (j) qui correspond à la somme en colonne (D_j). Ces deux (O_i , D_j) valeurs correspondent aux marges de la matrice (voir Figure 1-3). Cette propriété interne de la matrice qui permet d'opérer des sommes en lignes (ou en

²⁷ Les *matrices d'adjacence*, objet d'étude de la théorie des graphes, sont surtout analysées dans le champ de l'analyse des réseaux sociaux, des relations binaires entre les acteurs (*Social Network Analysis*).

colonnes) s'applique aussi aux marges. Le volume de flux, observé depuis un lieu (i) quelconque, noté (V_i), correspond alors à la somme de ses marges ($O_i + D_j$), quand le bilan net (S_i) correspond à leur différence. La Figure 1-3 présente le résultat de l'application de cette propriété d'additivité sur la matrice présentée dans la Figure 1-2, pour calculer le volume et le solde des flux, du point de vue des lieux. Par exemple, pour A, $V_i=(14+17)=31$ et $S_i=(17-14)=+3$.

Figure 1-3. Une matrice de flux et ses marges

		Fij				O _i	V _i	S _i
		A	B	C	D			
Marges : O _i : total des flux émis (somme en ligne) D _j : total des flux reçus (somme en colonne) V _i = O _i +D _j : volume S _i = D _j - O _i : bilan net* avec : i : lieux d'origine j : lieux de destination A: identifiants du lieu 8 : valeur du flux (A, B) de A vers B	A	-	8	5	1	14	31	+3
	B	6	-	3	2	11	25	+3
	C	10	4	-	1	15	26	-4
	D	1	2	3	-	6	10	-2
	D _j	17	14	11	4	46	92	0

* La formule est celle de la migration nette (entrées - sorties).

Nous n'irons pas plus loin dans la présentation des indices qui peuvent être dérivés de la propriété d'additivité des matrices de flux, car nous y reviendrons dans le chapitre 3. Examinons maintenant la valeur (F_{ij}) portée au croisement des lignes et des colonnes de la matrice. Située au cœur de la matrice, elle correspond à l'information échangée entre deux lieux sur laquelle nous portons notre attention dans cette thèse, compte tenu des difficultés de représentation cartographique qu'elle induit.

1.1.2 La valeur (F_{ij}) des cellules hors diagonale de la matrice

Au sens général, la valeur (F_{ij}) des cellules de la matrice de flux (F) correspond à une quantité mesurée qui s'est déplacée entre des lieux d'origine (i) et de destination (j), pendant un intervalle de temps $[t_0 - t_n]$, n correspondant ici à la seconde date. Sur la Figure 1-2 (et la Figure 1-3), la valeur du flux (A, B), c'est-à-dire transférée du lieu A vers le lieu B, est de '8' ($F_{ij}=8$ avec $i=A$ et $j=B$) ; inversement, celle du flux (B, A) est de '6' ($F_{ji}=6$ avec $j=B$ et $i=A$). La formulation des indices présentés pour le calcul des marges (O_i, D_j) de ces lieux peut ainsi être étendue aux valeurs de flux : le *volume bilatéral d'échanges* (V_{ij}), observé entre un lieu (i) et un lieu (j) correspond alors à ($F_{ij} + F_{ji}$) quand le *transfert net bilatéral d'échanges* (S_{ij}) correspond à ($F_{ji} - F_{ij}$).

Pour alléger la dénomination de ces deux notions relatives aux flux, de *volume bilatéral d'échanges* (V_{ij}) et de *transfert net bilatéral d'échanges* (S_{ij}), nous recourons à une terminologie abrégée qui correspond respectivement au *volume bilatéral* (V_{ij}) et au *solde bilatéral* (S_{ij}). Ces notions seront ainsi distinctes des termes que nous utiliserons pour les lieux à savoir : le *volume* (V_i) des flux aux lieux et le *bilan net ou solde* (S_i) des flux aux lieux, illustrés sur la Figure 1-3.

Ainsi, en revenant à notre exemple (voir Figure 1-2 et Figure 1-3), le *volume bilatéral* est de 14, pour le cas du flux (A, B) observé entre les lieux A et B. Il est calculé par la formule suivante :

$$\begin{aligned}
 F_{(A,B)} + F_{(B,A)} &= F_{(B,A)} + F_{(A,B)} \\
 &= 8+6 \\
 &= 14.
 \end{aligned}$$

La quantité qui s'est déplacée entre deux lieux, pendant une période, peut revêtir différentes acceptions en fonction de la thématique : le flux peut être migratoire, commercial ou encore financier, pour ne citer que ceux-là. Le choix de ces exemples n'est toutefois pas anodin puisqu'il s'agit là des trois grands types de flux couramment appréhendés à l'aide de matrices en géo-démographie et géographie économique, de même qu'en géographie des transports.

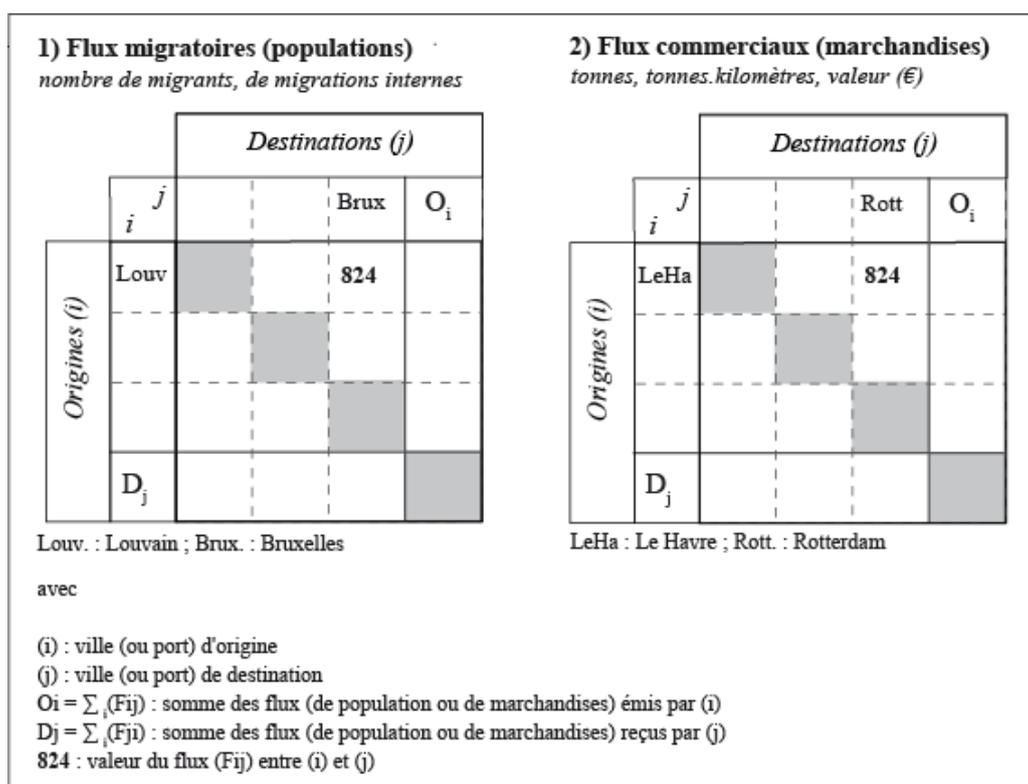
En démographie, le flux est migratoire et relié à un processus social qui intéresse l'individu tandis qu'en géographie économique ou des transports, le flux est commercial, financier ou de service, il s'intègre d'abord à un processus économique qui peut être appréhendé à travers la notion de trafic qu'il convient de distinguer de celle du flux. Le trafic est un type de flux particulier qui prend en compte le mode de transport (route, fer, mer, fleuve, air) et l'infrastructure technique organisée en réseau matériel (la voie routière ou ferroviaire, par exemple). Le flux que nous considérons ici est, à l'inverse, immatériel : il ne tient pas compte de l'infrastructure qui permet de mettre en relation deux lieux. Ce flux s'exprime entre un couple de lieux d'origine-destination (OD) et ne s'appuie pas sur un réseau à partir duquel il serait possible de reconstituer le cheminement exact de la quantité transportée, exprimée par la valeur du flux. C'est là un point important, caractéristique pour les matrices de flux qui nous intéressent : le cheminement entre un couple d'OD n'est *a priori* pas connu.

Le flux diffère aussi de la notion de déplacement qui suppose l'existence d'une trajectoire, d'un cheminement individuel, en démographie ou d'une chaîne, en géographie des transports. La trajectoire est composée de trajets ou de maillons auxquels sont associés des motifs de déplacement, des budgets-temps ou encore des modes de transport. Que l'approche géographique d'analyse d'une matrice, du moins celle que nous mettons en œuvre, soit d'inspiration démographique ou inscrite dans le champ de l'analyse socioéconomique des transports, sa structure est en apparence identique, tout comme la valeur du flux qu'elle présente. Pour le démontrer, nous allons considérer le cas de deux matrices de flux tout aussi semblables que distinctes, elles sont présentées dans la Figure 1-4, située page suivante.

La première matrice décrit des flux migratoires, de population humaine ; la seconde concerne des flux commerciaux de marchandises. Elles sont supposées présenter des valeurs dans toutes leurs cases mais, pour les besoins de la démonstration, nous n'avons mentionné qu'une seule valeur de flux représentée par la modalité 824.

Ainsi, peut-on lire sur la matrice des flux migratoires que 824 *migrants* se sont déplacés entre Louvain et Bruxelles, en Belgique, à moins qu'il ne s'agisse de 824 *migrations*, que l'on définit comme l'ensemble des déplacements des individus considérés comme des *migrants*.

Figure 1-4. Deux matrices de flux (migratoires et commerciaux)



Ces migrations sont internes²⁸ puisqu'elles ont eu lieu entre des villes situées à l'intérieur d'un même pays. Le pas de temps n'ayant pas été précisé, nous ne sommes pas en mesure de distinguer dans ces migrations internes s'il s'agit de navettes (déplacements quotidiens domicile-travail) ou de mobilités résidentielles, c'est-à-dire de transferts de lieu de résidence entre Louvain et Bruxelles. Les deux statistiques ne sont pas équivalentes et nous y revenons dans la section 1.3, page 39.

Sur la matrice de flux commerciaux, on observe que 824 marchandises se sont déplacées entre Le Havre, en France, et Rotterdam, aux Pays-Bas, à moins qu'il ne s'agisse de 824 colis indépendants ou groupés dans un même conteneur. En effet, les flux de marchandises entre Le Havre et Rotterdam s'effectuent majoritairement par voie maritime ou fluviale en conteneurs, selon un mode de transport par bateau. Cette matrice de flux commerciaux intéresse un phénomène qui s'est réalisé à l'échelle internationale sans toutefois franchir les limites de l'Union européenne, ce qui n'est pas sans conséquences sur la valeur que cette marchandise représente. Le fait que les lieux d'origine et de destination concernés soient des villes (des nœuds d'infrastructures routières, ferroviaires, portuaires, aéroportuaires, fluviales...) et parfois des portes d'entrée des pays concernés, suppose l'existence de postes douaniers ne serait-ce que pour dénombrer la quantité de biens en circulation, éventuellement la contrôler et la taxer²⁹. De ce fait, la matrice des flux commerciaux peut être interprétée en termes de volume, elle concerne alors 824 tonnes de produits – dont on ne sait pas s'ils sont tous identiques ou s'ils sont distincts –. Elle peut aussi être traduite en termes de volume par unité de distance et signifier 824 tonnes.kilomètres de marchandises (ou de produits). Cette même matrice peut, en outre, être

²⁸ Les *migrations internes* correspondent aux flux réalisés à l'intérieur d'une frontière nationale, entre des zones qui appartiennent à un même Etat. Ces flux sont, en fonction des auteurs et de leur thématique, qualifiés de flux nationaux ou encore de flux domestiques. Les *migrations internes* s'opposent aux *migrations internationales* réalisées entre différents états, nécessitant le franchissement d'au moins une frontière nationale.

²⁹ Les marchandises qui franchissent des limites nationales ont nécessairement fait l'objet d'une Déclaration d'Echanges de Biens (DEB) en douane.

interprétée en termes de nombre de paquets ou de colis ou, enfin, par rapport à la valeur que ces marchandises représentent : 824 K€ (kilos euros). Dans ce dernier cas, cette matrice serait alors une matrice de flux financiers, déguisée³⁰ en flux commerciaux, qui se sont produits entre le Havre et Rotterdam, pendant une période donnée.

Cette courte comparaison entre les valeurs de flux migratoires et celle de flux commerciaux illustre l'impossibilité de connaître *a priori* ce à quoi le flux correspond, en l'absence d'indications sur la manière dont il a été généré. Les deux matrices de flux présentées dans la Figure 1-4 sont identiques puisqu'elles possèdent la même valeur (824), mais en apparence seulement. Elles sont, en effet, fondamentalement différentes. Si l'on ne considère pas la référence temporelle, toutes les composantes caractéristiques des flux de ces deux matrices diffèrent : la thématique (flux migratoires *versus* flux de marchandises), la référence spatiale (flux interne *versus* flux internationaux) et surtout, l'unité de mesure des individus (des migrants, des migrations, des colis, des tonnages, des valeurs). Ces attributs propres à chaque flux, font référence au dispositif de collecte des données mis en place.

La signification des valeurs de flux archivées dans une matrice est, par conséquent, dépendante du cadre thématique, spatial et temporel dans lequel elle s'inscrit. En fait, la matrice de flux est une matrice de Berry (1964) particulière qui renseigne sur le transfert dans l'espace et dans le temps d'une certaine quantité (migratoire, commerciale ou financière).

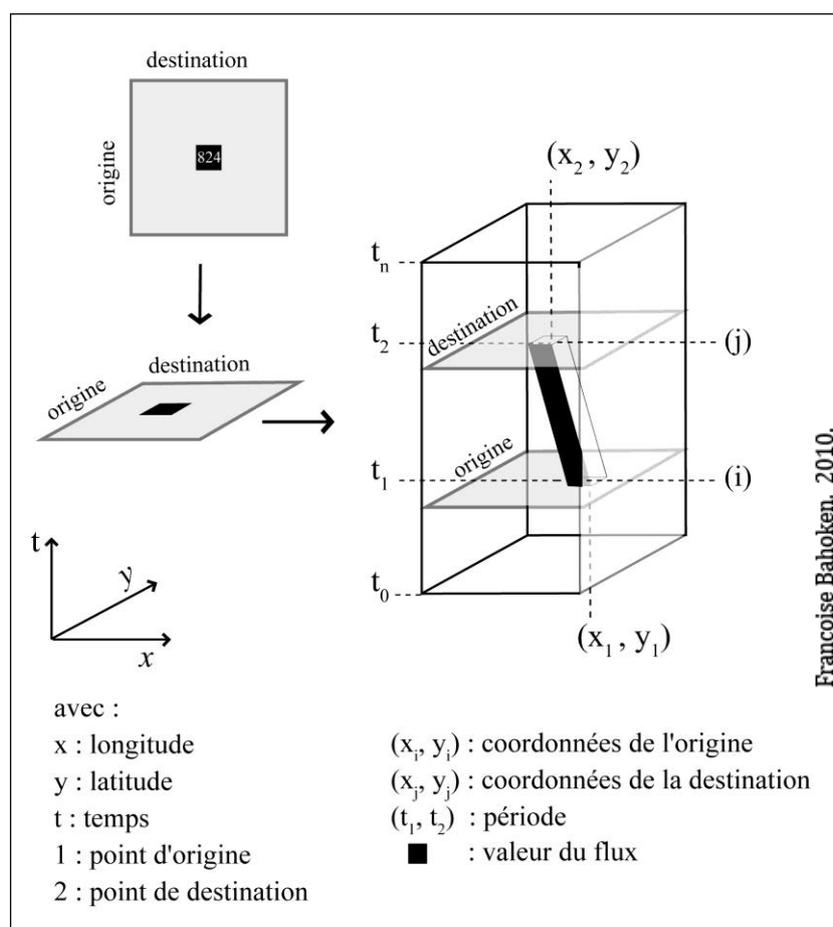
La Figure 1-5 située page suivante, décrit un parallélépipède qui illustre l'articulation des trois ou quatre (selon que la dimension spatiale en compte une ou deux) composantes (ou axes) qui concourent à la génération de la valeur du flux.

La thématique (z), en noir sur la figure, correspond à la quantité du type de flux dont il est question (de personnes, de biens) ; l'axe temporel (t) fait référence à la période d'observation du flux [$t_1 - t_2$] retenue pour la mesure ; la composante spatiale correspond à la résolution ou à l'échelle du maillage territorial utilisé pour la mesure, elle intéresse deux axes : l'axe des longitudes (x) et l'axe des latitudes (y) ; l'espace des échanges a été décomposé en deux espaces qui correspondent aux lieux d'origine et de destination.

En raison de cette nécessaire articulation de composantes qui caractérisent la mesure du flux, la réponse à la question de savoir ce que signifie précisément la valeur de la cellule (Fij) d'une matrice, au-delà d'une simple comparaison directe (migrants, marchandises ou villes/ports), réside dans le choix des caractéristiques de ces composantes, la période de temps par exemple.

³⁰ Cela n'est pas tout à fait exact puisque les flux financiers internationaux sont en réalité appréhendés par les montants alloués (par l'Union européenne, par exemple) au titre de l'aide au développement, par les investissements des firmes, par les Investissements Directs Etrangers (IDE) ou encore par les remises des migrants dans leur pays d'origine.

Figure 1-5. Définition du flux dans un aquarium d'Hägerstrand



La particularité de la mesure du flux est qu'elle résulte en pratique d'une double procédure de filtrage et d'agrégation spatio-temporelle mise en œuvre dans le cadre d'un dispositif de collecte nécessairement normalisé.

1.2 La mesure du flux : une double agrégation spatio-temporelle

Les valeurs disponibles dans les matrices de flux fournies par les organismes producteurs de données sont générées dans le cadre de dispositifs de collecte de données statistiques qui sont caractérisés par des spécifications technologiques, techniques, administratives et politiques. Les méthodes se caractérisent également par la sélection de critères sémantiques, propres à la population (statistique) dont il est question. La donnée collectée est dès lors toujours normée. Elle n'est de fait lisible et compréhensible que par rapport à la procédure de mesure appliquée dans le cadre du dispositif qui a permis sa génération. Les flux migratoires, par exemple, sont définis au regard de la définition des migrants retenue dans le cadre d'un recensement de population.

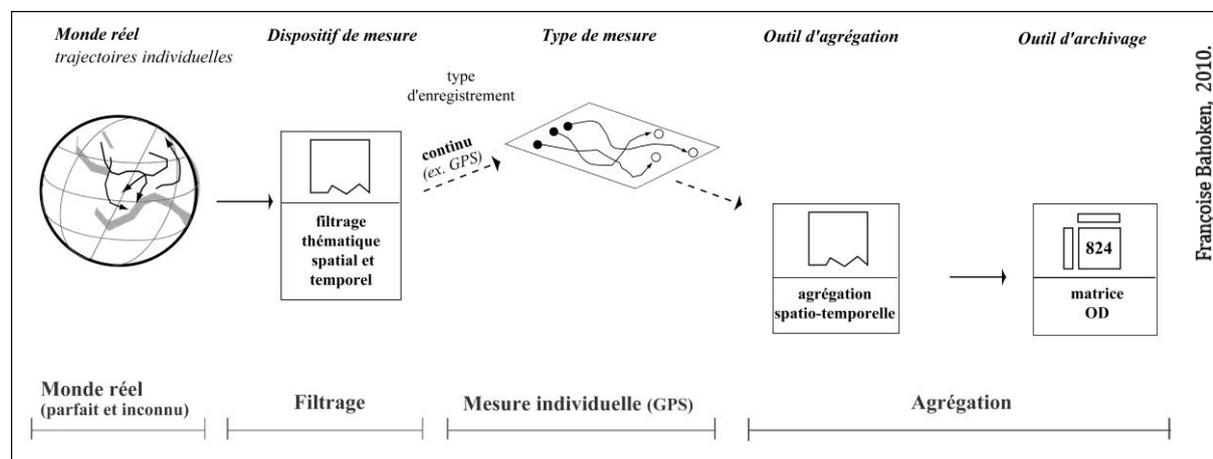
La particularité de la mesure des valeurs de flux est d'être menée dans le cadre d'une procédure appliquée à deux niveaux thématiques : au niveau élémentaire des déplacements des individus (statistique) ainsi qu'à celui de « paquets » d'individus. Le « paquet » étant, selon l'expression de P. Haggett (1973), l'ensemble formé par un groupe d'observations individuelles rassemblées, il correspond à la quantité élémentaire du flux qui n'est, en pratique, que rarement accessible. La valeur du flux résulte donc du dénombrement de déplacements individuels qui se sont toujours produits (ou qui vont l'être) pendant une période de temps et sur un espace délimités. Ces déplacements sont regroupés pour diverses raisons liées au dispositif de mesure.

En démographie, la mesure des flux migratoires est effectuée sur un (ou plusieurs) individus ou bien sur des événements : la naissance, le décès ou la migration sont les trois grands types d'événements analysés. La mesure est réalisée de manière transversale, c'est-à-dire à une date (t) précise ou sur une période $[t_1 - t_2]$; ou de manière longitudinale, au cours de la vie d'un (ou de plusieurs) individus ; depuis un événement particulier (sa naissance, par exemple) jusqu'au jour de l'enquête (ou à son décès, par exemple) et sur un espace défini par des zones d'origine et de destination du déplacement. La mesure des flux migratoires est par conséquent relative à un nombre de migrants ou à un nombre de déplacements (de leurs déplacements), les migrations.

Dans le champ des transports, cette même mesure des flux de population est effectuée par le biais de comptages « d'individus » réalisés en différents points d'un réseau (routier, par exemple). La mesure est alors celle d'un mouvement en un (ou plusieurs) nœud(s) du réseau, pendant une période $[t_1 - t_2]$, sans que soient nécessairement connues les zones d'origine et de destination. La collecte de ce mouvement consiste à dénombrer les individus (et/ou leur véhicule) qui franchissent un point particulier du réseau (ou un linéaire, comme dans le cas des Enquêtes Cordon³¹ routières, par exemple), dans un sens ou dans un autre, avec ou sans prise en compte – par le biais d'une question spécifique – de leurs lieux d'origine ou de destination. La mesure des flux de marchandises est également effectuée au regard d'une unité de mesure propre : la tonne-kilomètre, par exemple.

Deux grandes familles de méthodes de mesure des flux peuvent être mises en œuvre (voir Figure 1-6). La première est celle de l'enregistrement (pseudo) continu sur un temps plus ou moins réel des déplacements individuels. Elle permet la collecte de trajectoires individuelles spatio-temporelles à partir desquelles seront reconstitués des flux de zone à zone. La seconde famille passe par l'utilisation d'un bulletin de collecte des déplacements, sous une forme rétrospective ou prospective.

Figure 1-6. Illustration de la mesure du flux



La collecte de données réalisée lors d'enregistrements (pseudo) continus, à l'échelle individuelle des déplacements, est mise en œuvre à l'aide d'outils spécifiques qui sont présentés dans l'Encart 1-1.

³¹ L'enquête routière Cordon est un dispositif de collecte de données qui portent sur des déplacements par voie routière, réalisés en provenance ou à destination d'une zone délimitée au sol par un « cordon ».

Encart 1-1. La mesure en continu des déplacements individuels

Le cas le plus satisfaisant de mesure des flux est celui de l'enregistrement en continu de déplacements individuels réalisés à l'aide des outils de *Global Positioning System* (GPS) ou encore de traces de téléphones portables fournies par différents opérateurs de services mobiles. C'est à partir de l'enregistrement des différentes positions du GPS ou des signaux émis par le téléphone portable que sont reconstitués des tableaux de déplacements qui décrivent des traces individuelles.

Cette trace est une information collectée de manière quasi-continue. Elle correspond à une estimation de la trajectoire de l'individu dérivée de ses positions dans le temps et dans l'espace, ou de celle du mobile. L'outil de mesure sert alors de marqueur des déplacements. Les trajectoires individuelles font ensuite l'objet de procédures de filtrage spatio-temporel (Caceres et al., 2007 ; Olteanu-Raimond et al., 2011 ; Oltéanu-Raimond, Bahoken, 2013). Le signal collecté en continu nécessite, en effet, d'être rendu discret pour pouvoir être exploité. La composante temporelle est d'abord segmentée selon un pas de temps régulier ou irrégulier. A toute position temporelle est ensuite affectée une position dans l'espace en fonction de la vitesse de déplacement. Le signal peut également être collecté à deux instants précis de la trajectoire : la mesure porte alors sur un flux mesuré sur un intervalle temporel, il est collecté au point de destination. Enfin, le signal peut être collecté de manière régulière ou irrégulière qui correspond à N localisations spatio-temporelles : la mesure est alors celle d'un flux sur plusieurs intervalles temporels. Pour la mesure des déplacements de marchandises, il convient de mentionner en outre l'enquête Chargeur (1988, 2004), dite ECHO, qui collecte les trajets d'une sélection de 10 000 envois de marchandises particulières, émis depuis la France.

La trajectoire spatio-temporelle collectée décrit le cheminement du migrant pour les flux migratoires, ou la chaîne de transport pour les flux de marchandises. Cette trajectoire correspond à la succession des lieux pratiqués par l'individu, ou qui sont impliqués dans les opérations de transport d'un bien ; elle est formée par au moins un déplacement entre deux lieux (ou par au moins un trajet ou un maillon de la chaîne de transport). Ces données sont ensuite archivées dans des tableaux de déplacements individuels qui identifient les individus répondant aux critères de normalisation imposés par la méthode, c'est-à-dire ceux qui se sont déplacés sur la zone d'étude, pendant la période considérée. Ces tableaux de déplacements donnent ensuite lieu à une seconde procédure d'agrégation qui conduit à la génération d'une matrice de flux agrégée de lieu à lieu, à l'échelle des communes ou des pays, par exemple.

Cette mesure individuelle constitue un cas particulier d'information détaillée sur les composantes spatiales et temporelles du déplacement. Si elle est la plus satisfaisante, elle n'est que rarement disponible. La mesure des flux est, en général, réalisée dans le cadre d'enquêtes ou de recensements, à l'aide de bulletins ou de formulaires : c'est le cas, par exemple, pour la France, du recensement de population de l'Institut National de la Statistique et des Etudes Economiques (INSEE) qui fournit les fichiers MOBilités PROfessionnelles des individus (MOBPRO) sur les navettes intercommunales (les migrations domicile-travail) ; ou encore du fichier RAIL, réalisé par la Société Nationale des Chemins de Fer français (SNCF), qui fournit les flux de marchandises transportés par voie ferroviaire en France et en Europe.

La mesure réalisée au niveau élémentaire de l'individu, à l'aide de bulletins, entraîne la collecte d'une sélection de trajectoires individuelles qui correspondent à un plan de sondage, puisqu'il n'est pas possible de saisir l'exhaustivité des déplacements (Bavoux et al., 2005 : 51). La collecte est généralement rétrospective : pour les flux migratoires, l'enquêté est interrogé sur les déplacements

qu'il a réalisé sur une période donnée et pour les flux de marchandises, l'enquêté (qui correspond au chargeur ou à l'expéditeur) est interrogé sur le trajet et le produit qu'il a transporté. Dans ce dernier cas, la collecte peut également être prospective : c'est le cas de la mesure des flux internationaux de marchandises réalisée aux postes douaniers lors de laquelle l'expéditeur est interrogé sur le transport de marchandises qu'il s'apprête à effectuer (voir *infra*).

Une première procédure de filtrage qui articule l'ensemble des composantes (thématique, spatiale et temporelle) est également réalisée lors de la mesure (voir Figure 1-6). Celle-ci consiste à apposer à la mesure une grille ou un maillage dont la granularité est fonction du dispositif. Dans le cadre d'un recensement de population, entre autres, la composante thématique consiste à choisir la catégorie d'individus enquêtés : il s'agit, par exemple, de l'ensemble des individus de 15 ans et plus qui appartiennent à un même ménage et qui présentent, par ailleurs, certaines caractéristiques communes telle que celle d'occuper le même logement. La maille ou le pas de temps temporel peut correspondre à l'intervalle intercensitaire : les enquêtés sont alors interrogés sur les déplacements qu'ils ont réalisés depuis le précédent recensement ; la maille spatiale est alors celle de leur logement.

Cette mesure élémentaire est ensuite archivée dans des fichiers de détails. L'information spécifique aux déplacements est stockée dans un tableau spécifique qui porte sur les déplacements. Ce fichier fait l'objet, dans un premier temps, d'exploitations statistiques particulières avant de pouvoir être transmis aux utilisateurs.

Dans le cas des recensements de population, cette exploitation consiste à générer deux tableaux. Le premier est issu d'une exploitation dite principale, et comprend les données de l'ensemble des bulletins associées à une sélection de variables descriptives ; le second tableau, issu d'une exploitation dite secondaire, porte sur une sélection de bulletins associés à l'ensemble des variables caractéristiques des individus enquêtés.

Dans un second temps, ces différents tableaux font l'objet d'une procédure de généralisation³² qui prend la forme, dans notre cadre d'analyse, d'une agrégation spatiale de zone à zone. Les données collectées dans le cadre d'un recensement à l'échelle du logement sont ensuite agrégées par unité statistique de collecte, en Ilots Regroupés pour l'Information Statistique³³ (IRIS) pour les données issues de l'INSEE. Elles peuvent également être agrégées par commune ou par arrondissements municipaux, pour les grandes villes qui en possèdent. C'est cette seconde procédure qui conduit à la génération des valeurs disponibles dans la matrice de flux qui nous intéresse.

En raison de l'application de cette double procédure de filtrage et d'agrégation spatio-temporelle des données élémentaires qui portent sur des déplacements, l'information disponible dans une matrice de flux est, par essence, secondaire. Son degré de complétude dépend aussi des caractéristiques de la méthode de collecte appliquée au type de flux.

L'Annexe 1 présente les méthodes de collecte de l'information qui portent sur des flux, en prenant l'exemple de flux migratoires et de flux commerciaux.

Ces méthodes de collecte caractérisent ainsi la source des données à partir de laquelle sera générée la valeur des flux (migratoires, commerciaux, ...) archivés dans la matrice correspondante. C'est par conséquent à partir de l'identification des entités concernées (le migrant ou la marchandise), et de leur dénombrement (ou de celui de leurs déplacements : les migrations pour le migrant) que sont reconstitués les flux. Cependant, il n'est pas toujours évident de saisir le sens de ces entités élémentaires. Le principal problème de la collecte de données qui portent sur des mobilités spatialisées

³² Cette procédure de généralisation est indépendante des éventuelles opérations d'anonymisation ou permettant le respect de la confidentialité des données, réalisées par l'organisme fournisseur de données.

³³ Une définition de ce découpage territorial en IRIS est disponible sur le site de l'INSEE : <http://www.insee.fr/fr/methodes/default.asp?page=definitions/iris.htm> [Lien vérifié le 02/08/2015].

réside, en effet, dans la précision de la mesure d'un objet qu'il n'est pas toujours aisé de caractériser avec précision.

Pour les flux de marchandises la difficulté réside dans l'utilisation d'une unité de mesure exprimée en volume (tonnes, tonnes.kilomètre) ou en valeur, c'est-à-dire en prix courant qui doit intégrer des taux de change, en fonction d'un type de produit. L'objet « marchandise » poserait lui-même problème en raison de la multitude de produits existants et des différents lieux potentiels de sa production. Les marchandises sont, en effet, souvent générées dans différents lieux de production situés à l'étranger, en raison de l'internationalisation des processus de production et des questions liées à la sous-traitance internationale. Ainsi, quand bien même la marchandise serait étiquetée *Made in France*, la part des composants d'exportations, ou de celle des produits d'origine nationale, n'est pas toujours connue, ce qui conduit à une difficulté d'estimation de la valeur réelle des marchandises.

Pour les flux migratoires, c'est la notion même de migrant qui n'est pas toujours précise. Face à la difficulté d'identifier les migrants, au décompte de leurs déplacements dans le cadre spatiotemporel spécifique de l'observation des flux agrégés de zone à zone, nous avons développé une approche graphique compréhensive des différents concepts mobilisés pour leur mesure.

1.3 La construction de la matrice de flux : l'exemple des migrations humaines

Pour présenter la manière dont est générée une matrice de flux à partir de données collectées à l'échelle élémentaire d'un déplacement individuel, nous illustrons le cas général de la double procédure de filtrage et d'agrégation spatio-temporelle présentée sur la Figure 1-6 par un exemple, celui de flux migratoires collectés à l'aide d'un bulletin³⁴ de recensement. Nous choisissons cet exemple, car il permet d'illustrer parfaitement la procédure de construction d'une matrice de flux migratoires.

La collecte de données à l'aide d'un bulletin constitue, en effet, l'éventualité la plus plausible d'obtention de données collectées par un organisme tel que l'INSEE. Pour cela, nous partons de la définition démographique du migrant comme l'individu ayant effectué « [...] *au moins une migration, au cours d'une période donnée* » (Courgeau, 1973a : 95), c'est-à-dire au moins un déplacement pendant une période de temps $[t_1 - t_2]$ et sur un espace délimités. Le migrant est, en effet, « [...] *un individu dont le lieu de résidence est différent à la date du recensement* [...] » (Courgeau, 1975b : 28). Cette différence de localisation du lieu de résidence est estimée le jour de l'enquête, par rapport à une date antérieure puisque la donnée collectée lors d'un recensement est obtenue dans le cadre d'une méthode rétrospective. La mesure du flux migratoire porte par conséquent à la fois sur le migrant, l'individu qui s'est déplacé, que sur son nombre de déplacements, ses migrations. La matrice pouvant être une matrice de migrant ou de migrations.

Il est important de noter le fait suivant : pour une population donnée, le nombre de migrants et de migrations ne sont pas nécessairement des valeurs identiques. En effet, en analysant l'articulation des concepts de migrants et de migrations, plus précisément la différence entre ces deux mesures, M. Poulain a mis en évidence le fait que « [...] *le concept de migration peut être considéré comme un cas limite de celui du migrant si l'intervalle (t_0, t_1) tend vers 0* » (Poulain, 1981 : 15), c'est-à-dire si

³⁴ Des méthodes issues du champ des transports permettent également la génération de matrices de flux à partir de données ponctuelles de comptages réalisés en différents points d'un réseau, routier, par exemple. Elles consistent, pour l'essentiel, à créer un modèle de répartition des flux à partir des données ponctuelles collectées sur une période, puis à simuler des flux entre des zones d'origine et de destination précises. Ces méthodes sont bien décrites dans la littérature, notamment dans la thèse de S. Debaille, S., 1979, Reconstitution de matrices origine-destination en milieu urbain, *TEC*, n° 34/35, 6p. ; Chen, Y-S. 1993, *Utilisation conjointe d'un modèle de génération-distribution et des comptages de circulation pour la reconstitution d'une matrice origine-destination de trafic routier*. Thèse de Doctorat, ENPC, octobre 1993. Les méthodes que nous présentons ici ne sont pas issues de modèles de répartition du trafic, elles sont fondées sur une collecte directe ou indirecte de l'information auprès des individus.

l'intervalle considéré est très court. L'auteur ajoute, comme nous l'avions déjà noté précédemment, que « [...] *la mesure des migrations et celle des migrants sont essentiellement différentes : la première se rapporte à des événements, la seconde à des individus, la migration prend place à un moment déterminé, le migrant s'observe sur une période donnée* » (Poulain, 1981 : 16).

Face à la difficulté de saisir ces concepts et la manière dont ils interviennent dans la génération des valeurs des flux migratoires archivées dans une matrice (de migrants ou de migrations), l'approche graphique compréhensive des différents concepts impliqués illustre la génération d'une matrice de migrations.

1.3.1 Eléments sur l'approche compréhensive de la construction de la matrice de flux que nous proposons

La complexité d'articulation des concepts impliqués dans l'analyse quantitative des mobilités spatialisées nous a conduites à élaborer une approche graphique. Pour cela, nous avons recours à un exemple plausible de déplacement individuel : le récit de vie fictif d'une femme prénommée Marie. Nous considérons ce récit représentatif de la réalité d'une biographie individuelle reconstituée à partir d'un entretien semi-directif.

Nous envisageons l'approche par le récit de vie comme un outil pratique de documentation sur les concepts et les notions que nous souhaitons présenter. Cette approche suppose de comprendre les événements extraits d'un vécu individuel et de les repositionner dans le contexte de l'analyse spatialisée en géographie. Elle nous permet et c'est là son réel intérêt, de confronter la réalité des informations issues d'une retranscription subjective du vécu (l'histoire de la vie d'un individu) aux données quantitatives sur les flux dont on peut *in fine* disposer pour l'analyse, démontrant ainsi l'inévitable simplification des données empiriques par rapport à la réalité.

Le matériau que nous avons utilisé correspond au récit présenté dans l'Encart 1-2.

Encart 1-2. Extrait du récit de vie de Marie

« Marie est née le 1er Janvier 1960 à Romorantin (1), dans le Loir et Cher. Ses parents habitaient à l'époque La Goinfrerie, non loin de l'étang de Batarde. En 1970, ses parents divorcent. Marie se souvient qu'elle apprit cette nouvelle le jour de son anniversaire. Quelques semaines plus tard, Marie déménage avec sa mère, alors infirmière, à Tours (2). Marie obtient un baccalauréat A1, au lycée Gramont de Tours. Elle a alors 17 ans (3). Après trois années d'études universitaires, elle décide de se spécialiser en Etudes Africaines et, pour se faire, monte à Paris où elle arrive en 1980 (4). En 1985, Marie épouse Pierre alors qu'ils sont tous les deux doctorants. Pierre s'installe ensuite chez Marie (5). Après un doctorat obtenu en 1987, Marie peine à trouver le poste de Maître de conférence qu'elle obtient finalement en 1989, c'est-à-dire, la même année que Pierre (6). Marie est alors affectée à Metz tandis que Pierre l'est à Strasbourg. Le couple décide de s'installer à Metz. Pierre s'installe dès lors dans une existence ponctuée entre ses navettes hebdomadaires et ses déplacements professionnels en France, ou à l'étranger. (7) Le premier bébé de Marie naît en 1990, le second en 1992 et le troisième, en 1995. La naissance de cette troisième fille conduit Marie à déménager à Strasbourg, dans un appartement, puis dans une maison en 1998 qu'ils achètent (8). Mais en 2010, Marie et Pierre divorcent. Marie décide alors de retourner vivre à Paris (9). Elle obtient son poste de Professeure à 60 ans, en 2020 (10) [mais ce n'est pas bien grave car il lui reste encore dix ans de carrière !]. A la retraite à 70 ans, Marie s'installe à Romorantin. Sa mère décède quelques mois plus tard et elle hérite de la maison familiale (11). En 2040, Marie est foudroyée par un infarctus alors qu'elle vient d'intégrer la Maison de Retraite de Tours (12) ».

Les chiffres situés entre parenthèses renvoient à des situations sur lesquelles nous revenons dans le texte et au sein des figures pour situer le moment de la vie de l'individu dont il est question, reconstituer ses déplacements. La trajectoire biographique de Marie est marquée par ses différents déménagements (migration ou mobilité spatiale) et changements de situation sociale (mobilité sociale) ou familiale.

Nous allons reconstituer³⁵ les migrations de Marie à partir de l'analyse des données localisées de ce récit. Cela suppose l'identification des périodes lors desquelles Marie est migrante de celles d'immobilité³⁶, au regard d'un référentiel temporel et de la méthode de collecte mise en œuvre.

Pour représenter l'articulation et l'alternance entre les phases de mobilité – qui définissent le statut de migrant d'un individu – par opposition à celles d'immobilité qui correspondent au « non migrant », nous avons recours aux concepts développés dans le cadre de la *Time Geography* que nous appliquons à la trajectoire résidentielle de Marie. La Figure 1-7, située page suivante, présente l'ensemble des lieux de résidence fréquentés par Marie au cours de sa vie. Les chiffres mentionnés sont référencés dans l'Encart 1-2.

L'ensemble des lieux de résidence occupés par l'individu au cours de sa vie définit son *espace résidentiel* ou espace de vie. Au sens de la *Time Geography*, l'espace individuel des transferts de lieux de résidence d'un individu définit son *domaine résidentiel*, tandis que l'espace temporel définit son *prisme résidentiel*. Dans la version classique du modèle élaborée pour des déplacements quotidiens, la figure représente un *domaine* et un *prisme d'activité*. Ce dernier est défini par l'ensemble des lieux d'activités de l'intéressée. Le *prisme spatio-temporel* articule donc le domaine résidentiel et le prisme résidentiel. Il se définit graphiquement par un volume à l'intérieur duquel la trajectoire résidentielle de Marie (ou celle des activités quotidiennes d'un individu quelconque) se déploie. Il est délimité dans sa partie basse, par un plan en deux dimensions (x,y) qui correspond à l'espace résidentiel. Dans sa partie haute, le prisme résidentiel est délimité par l'axe vertical qui correspond au temps. Les *stations* (lignes verticales) indiquent les moments d'arrêt de la trajectoire.

Le prisme de Marie est présenté sur la Figure 1-7. Les stations renseignent sur les temps d'arrêt dans les lieux de résidence de Marie, par conséquent sur ses *durées de présence* dans chacun d'eux. Ces durées de présence sont représentées par la hauteur des lignes verticales. Ainsi, le statut d'immobile de Marie devient partie prenante du processus d'analyse de la mobilité résidentielle. Son intégration dans la figure lui rend une épaisseur temporelle : « *Même ce qui est immobile dans l'espace a une étendue dans le temps* » (Chardonnel, 2001 : 135). Le transfert³⁷ de résidence est, quant à lui, figuré par les lignes parallèles à l'axe des abscisses.

Des phases de cooccurrences sociales et/ou spatiales de Marie, avec d'autres membres de son ménage ou de sa famille auraient pu être mises en évidence sur cette Figure 1-7 (Bahoken, Gauvin, 2005). Par exemple, si nous avions représenté séparément la trajectoire résidentielle de Pierre, son mari, nous

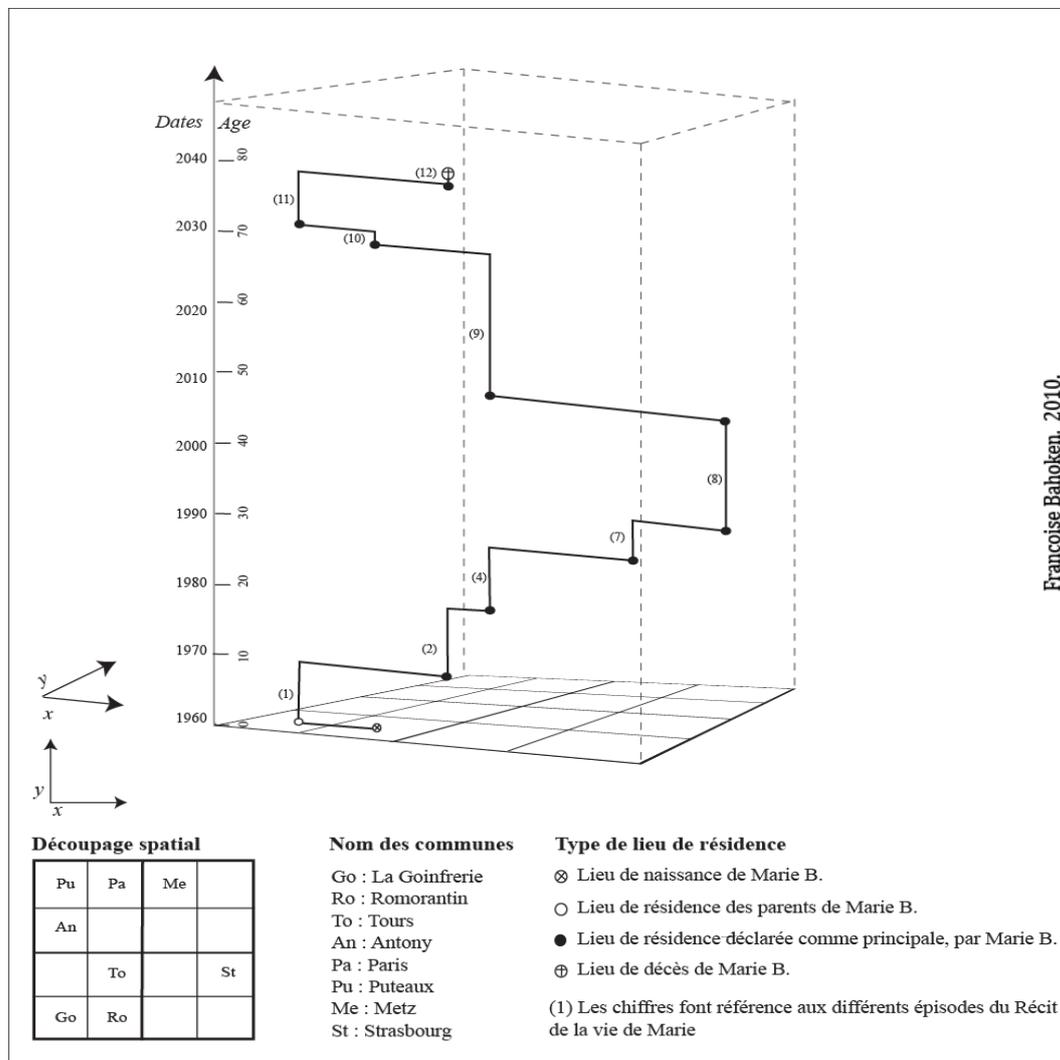
³⁵ Dans le cas des recensements de population, la collecte de l'information passe généralement par la réponse à une question unique sur la localisation antérieure (au jour de l'enquête) du lieu de résidence. Cette information retourne un effectif de migrants à deux dates, c'est-à-dire deux stocks et un transfert migratoire. A noter que c'est à partir de ces informations collectées auprès de l'ensemble des individus qui font l'objet de l'enquête que sont dénombrés le nombre de migrants.

³⁶ L'immobilité peut être entendue de deux manières : d'une part, comme la position de l'individu qui n'a pas réalisé de déplacement au cours de la période considérée et d'autre part, comme celle dont le déplacement n'a pas franchi de limite de zone.

³⁷ Pour les besoins de la démonstration, nous avons arrangé la formalisation classique du prisme en introduisant une double échelle temporelle, puis en qualifiant les différents lieux de résidence. Nous avons aussi fait le choix de ne pas introduire la notion de vitesse de déplacement : celle-ci l'aurait été par une inclinaison des pentes qui représentent les épisodes de déplacement. Cette notion peut, en effet, être considérée comme négligeable à l'échelle d'une trajectoire biographique.

aurions pu avoir l'opportunité d'analyser graphiquement leurs périodes conjointes de fréquentation de certains lieux (cooccurrence sociale des deux personnes) dont certaines correspondent à une période de co-résidence (cooccurrence spatiale dans le même lieu).

Figure 1-7. Le prisme spatio-temporel résidentiel de Marie



On observe ainsi sur la Figure 1-7 une variation du statut de Marie, selon les périodes de sa vie : celle-ci peut être sédentaire ou immobile sur certaines d'entre elles (qui correspondent aux traits verticaux), mobile et considérée comme migrante sur d'autres. Ce statut de l'individu peut aussi changer en fonction de la manière dont la question sur la localisation de sa résidence antérieure est posée dans le cadre d'un recensement, ce qui constitue un premier élément d'explication de l'incertitude de la mesure des flux migratoires.

Pour le comprendre, supposons que l'enquêteur du recensement de 1965 pose à Marie les deux questions [Q1] et [Q2] suivantes :

– « Où habitez-vous le 1^{er} janvier 1960 ? » [Q1]

– « A Romorantin » répond l'enquêtée.

– « Quand et où êtes-vous née ? » [Q2]

– « le 1er Janvier 1960, à Romorantin » répond l'enquêtée.

Dans le recensement de 1965, l'enquêteur sait que Marie réside à la Goinfrerie puisque le recensement à lieu sur son lieu de domicile. Cependant, la réponse de *Marie* à la question [Q1] est inexacte puisqu'elle est née le 1^{er} janvier 1960 et qu'à cette date, elle était présumée résidente à Romorantin, dans le Loir-et-Cher (41), là où est située la maternité de l'Hôpital local de Romorantin-Lanthenay. De plus, à la même date, ses parents résidaient à la Goinfrerie, une petite commune située à proximité de Romorantin (41194) mais qui dépend de la commune de Pruniers-en-Sologne (41185). Ainsi, du point de vue du recensement, la réponse de Marie à la question de la localisation antérieure (en 1960) de sa résidence est ici erronée. Le découpage territorial utilisé pour la mesure est, en effet, communal : les communes sont identifiées par leur code INSEE et non par leur code postal, par exemple. Marie aurait dès lors dû déclarer résider à la Goinfrerie, réponse à la question [Q1], car elle est administrativement rattachée à la résidence principale de ses parents. A noter que si l'identifiant communal retenu pour la mesure est le code postal, la réponse de Marie à la [Q1] est par simple hasard exacte puisque les communes de Pruniers-en-Sologne et de Romorantin ont le même code postal (41200). Cette erreur est une manifestation de *l'effet-mémoire* de l'enquête qui consiste à confondre ou à se méprendre sur les dates de survenue des événements qui sont anciens ou peu importants pour lui.

Le statut de l'individu diffère aussi si l'on s'en tient à la déclaration, exacte, de ses parents vu que dans ce cas, Marie n'est pas comptabilisée comme migrante. Le Tableau 1-1 présente la variation du statut de migrante de Marie, en fonction de sa réponse à la question de la localisation de son lieu de résidence à une date antérieure à celle du recensement.

Tableau 1-1. La variation du statut de migrante de Marie en fonction des réponses

	Effectif	Déclaré par Marie [Q1]	Déclaré par un parent [Q1]
Résidence en 1960	Romorantin	Romorantin	La Goinfrerie/Pruniers
Résidence en 1965	La Goinfrerie/Pruniers	La Goinfrerie/Pruniers	La Goinfrerie/Pruniers
Statut	migrante	migrante	immobile

La réponse de *Marie* à la [Q2] est en revanche exacte : elle est bien née à Romorantin, en 1960 : son *lieu de naissance* en 1960 et son *lieu de résidence principale* en 1965 sont différents. Sur la base de la réponse à [Q2], Marie est donc comptabilisée dans le stock de migrants et il lui est attribué une seule migration interne. Son changement de résidence a lieu pendant la période de référence, en 1960, à environ trois jours d'intervalle.

Sur la Figure 1-7, nous n'avons représenté que les changements de lieu de résidence de Marie, ce qui nous a permis d'identifier les périodes de mobilité et d'immobilité ; la mobilité n'ayant été définie qu'en fonction des changements de localisation des lieux de résidence, elle est appréciée au regard du dispositif de collecte considéré, comme l'a montré l'exemple du recensement de 1965. Cependant, un individu peut avoir effectué des déplacements depuis le lieu déclaré comme résidence principale et d'autres lieux dans lesquels il a également résidé mais qui ne sont pas considérés comme sa résidence principale au sens administratif. Il s'agit, là encore, du résultat de l'application d'une grille de collecte normée des données. Ces autres lieux de résidence peuvent avoir été fréquentés pour des raisons familiales (résidence secondaire, résidence familiale), de loisir (résidence de tourisme) ou encore dans un cadre professionnel (logement de passage, hôtel). La distinction entre le lieu de résidence déclarée comme principale et les autres lieux de résidence de l'individu, au cours d'une période donnée et sur un espace délimité, est fondée sur un indicateur spatiotemporel : du temps passé et/ou de la fréquence de passage dans un lieu. Un lieu sera ainsi intégré au *domaine résidentiel* de l'individu s'il y a séjourné plus de n nuits cumulées dans l'année (n étant défini par l'organisme enquêteur). L'introduction d'une durée de présence conduit à considérer d'autres lieux de résidence, dans certains

cas particuliers à changer le statut du lieu de résidence et, à le requalifier, et inversement pour les autres lieux.

La réponse à une question unique posée dans le cadre d'un recensement ne permet de collecter qu'un seul transfert du lieu de résidence, c'est-à-dire qu'une seule migration pour l'ensemble des individus considérés comme migrants au cours de la période de référence. Or, l'on sait que le nombre de migrations varie en fonction des périodes de la vie (Courgeau, 1987). Cette incertitude sur la mesure des migrants caractérise le *filtre thématique*, elle se répercute sur le nombre de migrations collectées en raison de l'application de *filtres spatio-temporels* lors de la mesure individuelle.

1.3.2 La procédure de filtrage de la mesure

La résolution des filtres spatio-temporels appliqués sur la mesure du nombre de migrants et de leurs migrations est essentielle car elle définit les valeurs (i,j) portées *in fine* sur la matrice, celles qui feront l'objet de l'analyse. Cette résolution varie selon les méthodes de collecte. La différence entre ces deux mesures individuelles, sur une période donnée est, en effet, imputable aux normes administratives, aux découpages temporels et administratifs utilisés lors de la collecte des données. Ces découpages opèrent tels des filtres de la mesure individuelle dont l'application a pour conséquence qu'en fonction du pas de temps et de la résolution spatiale, le déplacement sera comptabilisé ou pas. Ceci explique le caractère par essence lacunaire des données qui portent sur des flux.

« [...] *contrairement aux autres phénomènes démographiques, les migrations ne sont presque jamais comptées intégralement, soit qu'on perde des migrations multiples, ou les deux migrations en sens inverse, qui ont pu avoir lieu au cours de la période, soit qu'on omette les changements de résidence à l'intérieur des unités territoriales, dont on étudie les échanges par migration* ». (Courgeau, 1973b : 511).

Dans le cas des recensements de population, la dépendance de la statistique du nombre de migrations collectées en fonction de l'intervalle temporel considéré est une manifestation de *l'effet du filtre temporel*. C'est par conséquent à partir des données collectées à l'échelle des individus qui ont déclaré avoir changé de résidence entre deux dates précises que sont reconstitués les événements qui portent sur leurs déplacements, en l'occurrence les migrations. Le nombre de migrations d'un individu est donc d'abord filtré selon une référence temporelle (c'est le filtrage temporel) puis spatiale (c'est le filtre spatial ou, plus précisément territorial : nous verrons dans la section 1.3.3.1, située page 48 que la mesure du flux est liée au franchissement d'une limite de zone). Le nombre de migrations est donc directement dépendant de l'effectif de migrants, mais les deux valeurs ne sont pas équivalentes.

1.3.2.1 Le filtre temporel

Le filtre temporel est le découpage du temps utilisé pour la collecte des données sur les mobilités : plus le pas de temps sera grand, moins la mesure sera précise. Cet effet peut être expliqué d'une autre manière : plus l'individu aura effectué de migrations au cours d'une période donnée, plus l'écart entre sa migration et son nombre de transferts de résidence sera important, et plus les données collectées seront incomplètes. Ce cas de figure peut être illustré par les périodes (7) et (8) de la vie de Marie durant lesquelles elle a occupé trois logements (un à Metz et deux à Strasbourg) et effectué deux migrations internes. Comme précédemment, supposons la réponse à la question qui lui serait posée, dans le cadre du dernier recensement exhaustif français de 1999.

Dans le recensement de 1999, l'enquêteur sait que Marie réside dans une maison à Strasbourg, puisque le recensement a lieu à son lieu de domicile. Il lui pose alors la question suivante [Q3] :

– « Où habitez-vous le 1er Janvier 1990 ? » [Q3]

– [Variante : Où habitez-vous au précédent Recensement de la Population (RP) ?]

– « A Metz » répond l'enquêtée.

La réponse de Marie est exacte : elle est installée à Metz avec Pierre depuis 1989, date de leurs recrutements professionnels respectifs. D'après la réponse de Marie à la question [Q3], l'enquêteur dispose des informations suivantes :

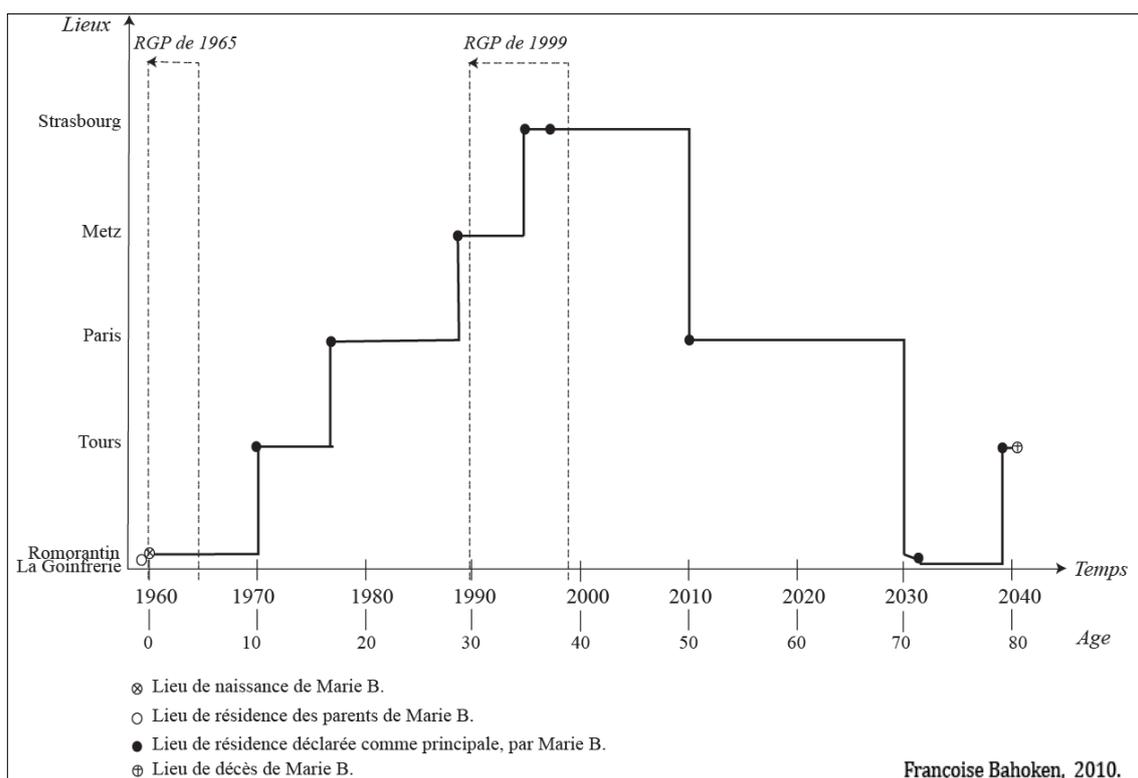
- le lieu de résidence de Marie en 1999 : Strasbourg ;
- le lieu de résidence de Marie en 1990, tel qu'elle le déclare : Metz ;
- le nombre de migrations enregistrées, dans le cadre du RGP, entre 1990 et 1999 : 1.

Etant donné que nous avons accès au récit de vie de Marie, (voir Encart 1-2), nous disposons d'une information supplémentaire :

- le nombre de migrations réelles entre 1990 et 1999 : 2.

D'après sa déclaration à la [Q3], Marie est comptabilisée comme *migrante* entre 1990 et 1999 et une seule de ses migrations internes est enregistrée dans les statistiques du recensement. Pourtant, elle a effectué deux *migrations internes* : entre Metz (I_1) et l'appartement (I_2) de Strasbourg qu'elle occupe en 1995, après la naissance de sa troisième fille puis, entre l'appartement (I_2) et la maison (I_3) située elle aussi à Strasbourg. La Figure 1-8 illustre les intervalles temporels considérés par les recensements de 1965 et 1999 pris en compte pour la mesure du nombre de migrations de Marie.

Figure 1-8. Le décompte des migrations de Marie, selon différents intervalles temporels



Le *migrant* se définit ainsi comme l'individu qui *déclare* un changement du lieu de sa résidence principale entre deux dates précises, quelle que soit l'ampleur, la durée et le nombre de ses déplacements pendant la période. Le filtre temporel est donc une illustration du rôle joué par le découpage temporel utilisé pour la mesure du nombre de migrants et, par déduction, de celle du nombre de migrations.

La procédure de filtrage temporel engendre l'exclusion de deux catégories de migrants : la première est celle de ceux qui ont changé de résidence dans un pas de temps plus fin que celui défini pour la collecte. C'est le cas de la réponse de Marie à la [Q3]. La seconde catégorie de migrants exclus concerne les individus qui sont décédés entre temps. Ainsi, en raison de l'effet du filtre temporel, les données collectées sur le nombre de migrants peuvent être au mieux qualifiées de « *migrants apparents* » (Royer, 2009 : 11).

Aux côtés de l'effet du découpage temporel lié à la méthode de collecte mise en œuvre, il convient de noter deux autres effets qui influent sur le statut de migrant d'un individu et, par conséquent, sur le nombre de migrations qu'il a pu effectuer sur une période donnée. Ces effets sont liés à la composante spatiale de la mesure individuelle.

1.3.2.2 *Le filtre territorial*

Dans la réponse à une question du recensement de la forme : « *Où habitez-vous il y a cinq ans ?* » se trouve intégrée la référence au lieu de résidence. En pratique, les données qui portent sur les lieux sont détaillées à l'adresse du logement près. Toutefois, elles ne sont transmises par les organismes fournisseurs de données que sous une forme agrégée dans l'espace pour des questions de confidentialité, en France généralement à l'échelle des IRIS, depuis 1999. Quoi qu'il en soit, le nombre de migrations d'un individu dépend de l'échelon spatial considéré, du découpage territorial utilisé pour la mesure (Courgeau, 1973b). Si les changements de lieux de résidence sont collectés à l'échelle du logement, alors un même individu aura très certainement effectué plus de migrations que si elles sont collectées à l'échelle du département ou de la région : la résolution spatiale au niveau du logement est, en effet, beaucoup plus fine que celle du département ou de la région. Cette variation du nombre de migrations en fonction de l'échelle spatiale de collecte des données est une manifestation de *l'effet du filtre territorial de la mesure*.

Le filtre spatial est le découpage du territoire utilisé pour agréger les données sur les mobilités : il s'agit du maillage territorial de référence dont l'origine est de nature technique, ou politico-administrative. Le découpage communal a, par exemple, été défini dans un objectif de contrôle. Ce référentiel géographique intervient, en pratique, de deux grandes manières dans la mesure du flux. Le premier effet du découpage territorial est lié à l'imprécision³⁸ caractéristique des données qui portent sur les lieux issues de la retranscription du récit de la vie de Marie, même si l'on peut supposer qu'il a été collecté dans le cadre d'un entretien semi-directif réel. Celui-ci n'est pas totalement précis quant aux lieux mentionnés. Il est, en effet, indiqué dans le récit que Marie « *monte à Paris* » dès 1980, puis aucun changement de commune de résidence n'est mentionné jusqu'en 1989. Or, il n'en est rien puisque Paris signifie dans ce cas précis, la région parisienne, ce découpage territorial supra communal qui englobe la ville de Paris et quelques 1300 communes situées dans un voisinage d'une cinquantaine de kilomètres. En réalité, Marie occupera plusieurs communes de la région parisienne lors de l'épisode (4) du récit de vie. Pour le mettre en évidence, supposons que le détail de cette période est ainsi rapporté par sa troisième fille (voir Encart 1-3).

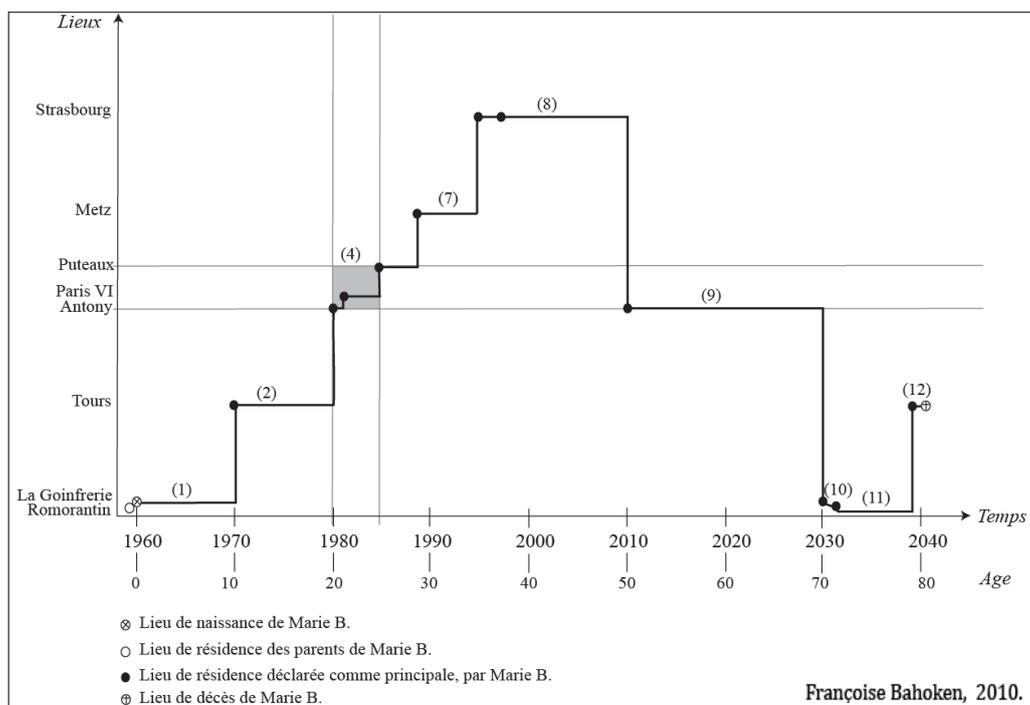
Encart 1-3. La période parisienne de Marie

« *Lorsque ma mère arrive à Paris en 1980, pour y faire ses études africaines, elle est logée dans un premier temps à la Cité Universitaire d'Antony où elle restera un an. Elle louera ensuite une chambre de bonne au 134 du Boulevard Saint-Germain, à Paris, pendant trois années. Elle s'installera ensuite dans un deux-pièces à Puteaux, où Pierre la rejoindra ensuite en 1985, après leur mariage [...]* »

³⁸ Cette imprécision des données est, en pratique, liée au niveau d'observation du phénomène, à sa résolution spatiale.

Ainsi, outre l'imprécision du récit sur les lieux mentionnés, le second effet du découpage territorial qui influe sur la mesure individuelle (sur le statut de migrant ou d'immobile) est lié à son échelle, à sa résolution. Dans la réalité, Marie aura, en effet, occupé trois logements successifs, lorsqu'elle est montée à Paris (voir Encart 1-3), comme l'illustre la Figure 1-9.

Figure 1-9. Le décompte des migrations de Marie selon le découpage territorial



Un seul des logements occupés par Marie est réellement situé dans la commune de Paris (75), les deux autres sont situés à Antony et à Puteaux, dans les Hauts-de-Seine (92). Sans tenir compte de la composante temporelle, on observe une variation du nombre de migrations de Marie en fonction de l'échelle spatiale considérée. Si ses migrations internes sont collectées à l'échelle régionale, alors Marie aura fréquenté tour à tour cinq régions (à noter qu'après la fusion des régions décidées en 2015, le nombre de régions fréquentées par Marie serait modifié rétrospectivement et porté à quatre). Marie aura effectué quatre transferts de résidence : la région Centre entre 1960 et 1980, la région parisienne entre 1980 et 1989, puis la Lorraine entre 1989 et 2010. Elle aura ensuite effectué une migration de retour en région Ile-de-France entre 2010 et 2030, avant de retourner en région Centre finir ses jours. En revanche, si les migrations internes de Marie sont collectées à l'échelle la plus fine qui est celle du logement, alors Marie aura transféré quatorze fois sa résidence principale, ce qui correspond à treize migrations résidentielles. Pour la seule région parisienne, Marie aura effectué deux migrations internes à la région entre 1980 et 1990. Cette variation dans le nombre de migrations en fonction de l'échelle spatiale de collecte des données est une manifestation caractéristique de *l'effet du filtre spatial de la mesure* individuelle qui conduit au décompte des migrants et de leurs migrations.

Le filtre spatial est donc lié au découpage du territoire utilisé pour la collecte des données individuelles, il correspond au maillage territorial de référence : plus la maille sera lâche et moins le nombre de migrations collectées sera important, comme l'indique D. Courgeau (1973b). Dans un tel cas, les données sont dites agrégées, généralisées ou sous-échantillonnées. A l'inverse, si le découpage territorial est trop fin, certains déplacements peuvent être exagérément isolés, ce qui peut conduire à un sur-échantillonnage.

On observe ainsi une variation du nombre de migrations en fonction des découpages temporels (voir

Figure 1-8) et territoriaux (voir Figure 1-9) utilisés pour la mesure. En pratique, plusieurs échelles d'observations peuvent être mobilisées : celle des communes, des départements ou encore des régions françaises, par exemple. Le passage de la mesure collectée à l'échelle des individus à une mesure agrégée à l'échelle du département s'effectue dans le cadre d'une procédure d'agrégation de ces « paquets » formés par l'ensemble des déplacements individuels.

1.3.3 La procédure d'agrégation spatiale des « paquets »

La procédure d'agrégation spatiale des déplacements individuels rassemblés par « paquets » caractérise l'effet du choix du niveau de résolution du maillage territorial sur la mesure, le découpage temporel ayant été sélectionné à l'étape antérieure, comme illustré par la Figure 1-6 (page 36). Elle consiste, en pratique, à transférer des données numériques depuis un niveau élémentaire (celui des IRIS, par exemple) vers un niveau territorial supérieur. Ce transfert des données d'un échelon à un autre entraîne un ré-échantillonnage vertical des valeurs qui correspondent à la multitude de petits événements individuels, observés de manière dynamique et mesurés de façon statique à l'aide d'un bulletin. C'est la raison pour laquelle les valeurs d'une matrice de flux qui correspondent à des mouvements ne sont jamais instantanées puisque les transferts de niveaux s'opèrent toujours dans un intervalle de temps borné $[t_0 ; t_1]$ et sur un espace d'étude délimité.

Ainsi, une information géographique disponible à une échelle d'agrégation importante (les régions, par exemple) est supposée avoir été ré-échantillonnée ou agrégée plusieurs fois. De ce fait, les données sont implicitement disponibles à plusieurs niveaux, surtout si ces derniers sont emboîtés, comme c'est le cas pour les découpages administratifs territoriaux en communes, départements et régions français. Cette disponibilité multi-échelle engendre un problème supplémentaire qui est celui de *l'effet de récurrence* (Grasland, 1997) qui consiste à dénombrer plusieurs fois un même événement, mais avec un niveau de généralisation variable.

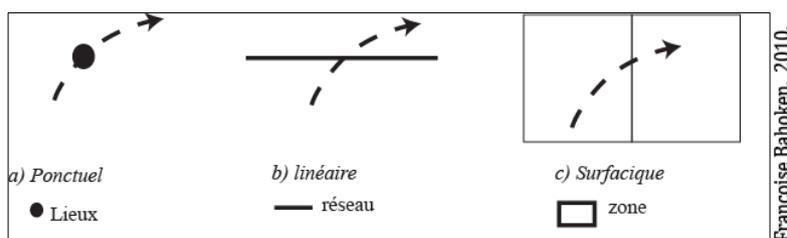
Sur le plan méthodologique, la procédure d'agrégation conduit ainsi à la mise en œuvre de deux processus qui visent à redécouper la trajectoire spatio-temporelle puis à fusionner les entités du niveau élémentaire en entités de niveau supérieur. La fusion concerne les critères liés à l'agrégation des différentes zones élémentaires qui seront transférées dans une même « région » de niveau supérieur. Les données de flux étant quantitatives, la fusion porte sur le choix de la fonction mathématique de fusion parmi lesquelles le dénombrement ou comptage des déplacements, la somme arithmétique, le minimum, le maximum ou encore la moyenne. Le second paramètre est d'ordre géographique, il est fondé sur un critère d'appartenance territoriale (les données disponibles à l'échelle des communes d'un même département seront agrégées) ou bien sur un critère de proximité. Cette proximité est soit continue et concerne alors une métrique, la distance kilométrique par exemple ; soit topologique ou discrète, elle fait alors référence à la notion de franchissement des limites de zones.

La procédure d'agrégation spatiale des flux s'appuie sur une fonction mathématique, généralement la somme arithmétique, ainsi que sur un critère de proximité topologique. Elle consiste à dénombrer puis à regrouper (à sommer) les déplacements individuels qui ont franchi un point particulier de l'espace d'étude au cours d'une période. Une telle mesure s'apparente à celle d'un débit mais, dans notre cadre d'analyse, la vitesse n'est pas prise en compte.

1.3.3.1 Le franchissement de la limite de zone

Dans notre cadre d'analyse des flux de lieu à lieu, la mesure du débit doit être entendue comme celle du franchissement d'une limite de zone géographique, présenté sur la Figure 1-10.

Figure 1-10. Illustration du franchissement d'une limite de zone

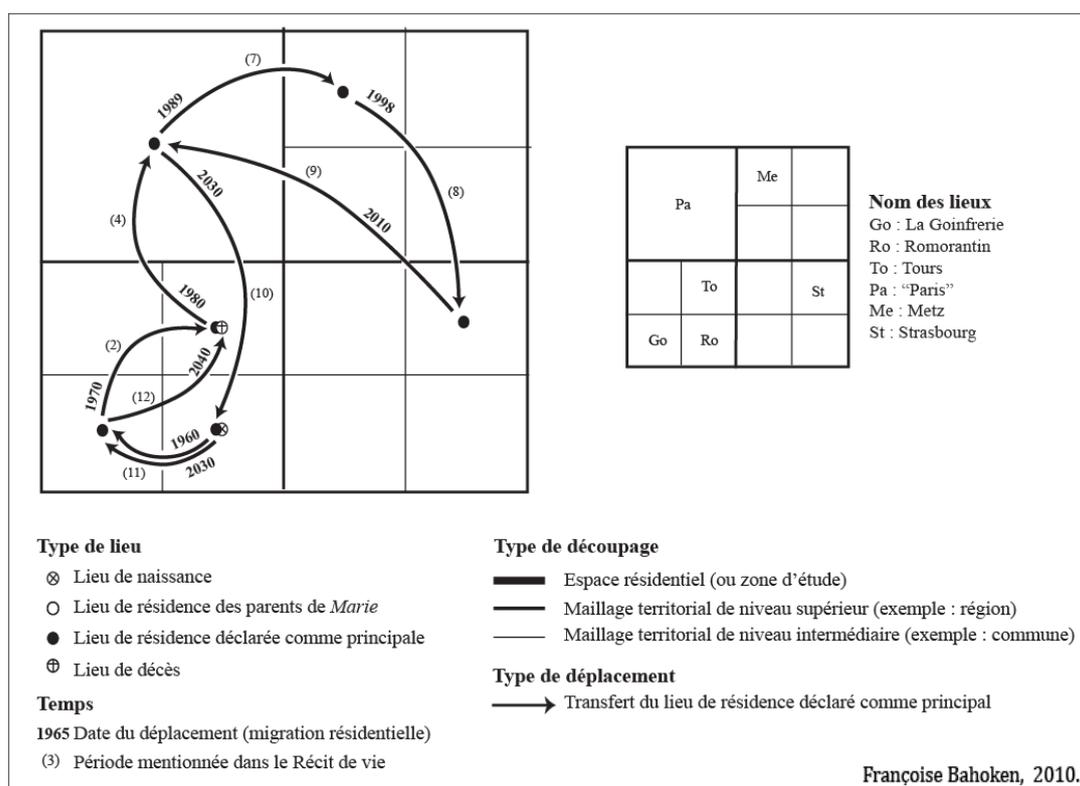


La limite de zone est une notion fondamentale dans l'analyse et la cartographie des flux puisque les mesures de l'effectif de migrants et du nombre de leurs migrations s'appuient, d'une part, sur son franchissement et d'autre part, sur la résolution de la zone.

Pour saisir les effets du choix du maillage sur le décompte du nombre de migrations d'un individu, revenons à l'épisode (4) qui correspond à la vie parisienne de Marie et changeons de modèle graphique.

La Figure 1-11 illustre la trajectoire spatiale résidentielle de Marie sur un plan en deux dimensions et à deux échelles, correspondent à une partition territoriale en deux échelons : celui des communes mentionnées dans le récit et celui des régions.

Figure 1-11. La trajectoire spatiale de Marie dans le plan

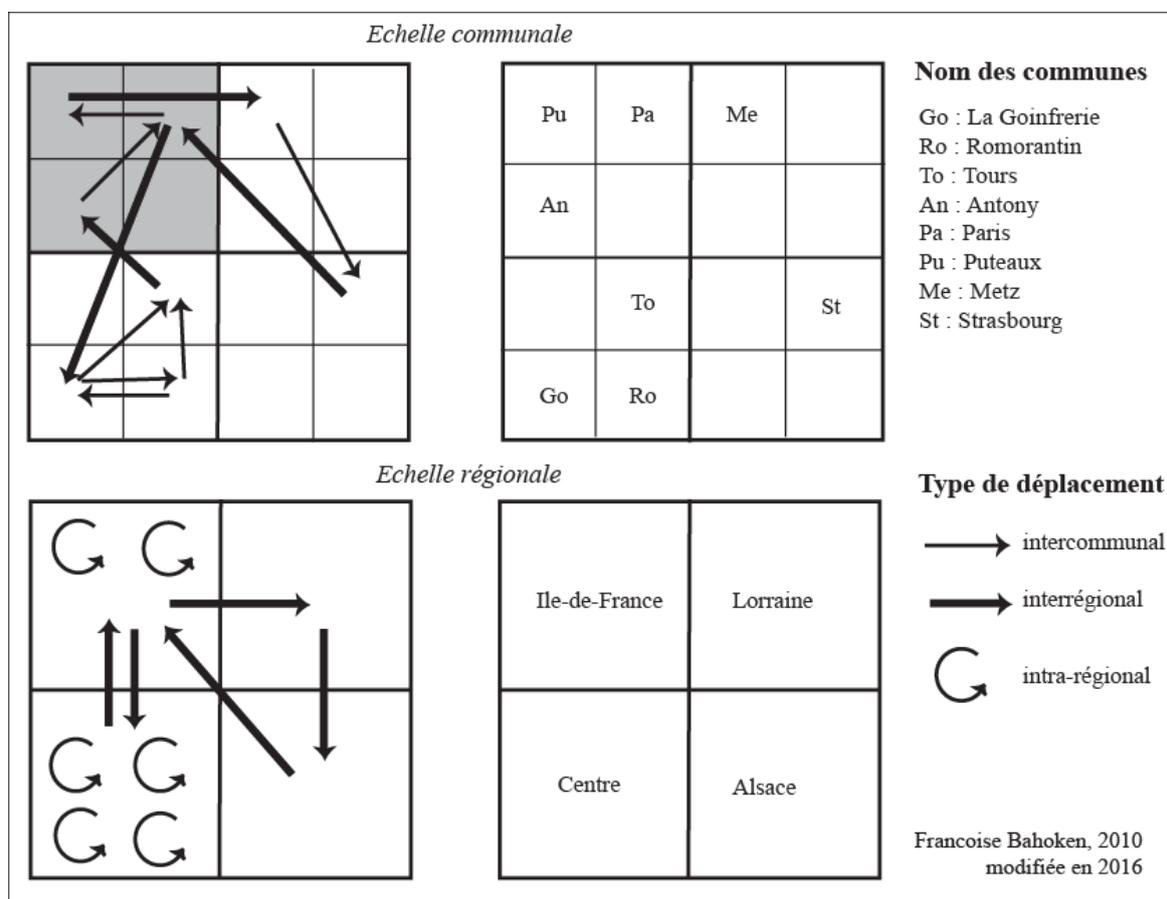


Pour illustrer l'effet du franchissement de limites de zones qui permet le décompte du nombre de migrations de Marie, c'est-à-dire l'effet de la procédure d'agrégation autorisant le passage de ses déplacements individuels à celui de l'agrégat, supposons *i*) que le périmètre de collecte de ses migrations soit constant et limité aux migrations internes et *ii*) que la mesure porte sur la période

1960-2040. Le choix de cette résolution temporelle nous permet de considérer l'ensemble de ses déplacements.

La Figure 1-12 est réalisée en s'inspirant de la représentation proposée par (Pumain et Saint-Julien, 2001 : 14) pour illustrer la procédure d'agrégation des déplacements de Marie présentés précédemment (voir Figure 1-11).

Figure 1-12. L'effet du franchissement des limites de zones



La partie haute de la Figure 1-12 représente les transferts de résidence intercommunaux de Marie à l'échelle des communes, la partie basse correspond aux déplacements inter et intra régionaux.

Etant donné que l'analyse de cette figure va conduire *in fine* à passer d'une information qualitative (un type de déplacement) à une information quantitative (un nombre de déplacements), les flèches qui illustrent le déplacement dans le plan sont désormais représentées par des lignes droites et non plus courbes. Aussi, pour faciliter la lecture, l'épaisseur du vecteur de déplacement entre deux lieux correspond à l'échelle spatiale à laquelle il est mesuré : les lignes épaisses concernent les migrations régionales, c'est-à-dire les déplacements interrégionaux, *externes* ou *interzones*, tandis que les lignes fines portent sur des déplacements intercommunaux, ces derniers sont, par conséquent, infrarégionaux, *internes* ou *intra zones*.

D'après la Figure 1-12, à l'échelle régionale, Marie n'est pas considérée comme migrante entre 1980 et 1990. Pourtant elle a bien migré entre les cinq régions, comme nous l'avons déjà mentionné (page 47). Elle a aussi effectué deux migrations internes au cours de la période considérée qui ne sont pas prises en compte dans le recensement. L'explication tient au fait qu'elle est sortie du territoire de la région parisienne au cours de la période de collecte : Marie résidait en 1980 à Antony mais en 1990, elle ne résidait plus en région parisienne, mais en Lorraine (voir Figure 1-9, page 47). Par ailleurs, les

deux transferts de résidence que Marie a réalisés en région parisienne, entre Antony, Paris et Puteaux, ne sont pas comptabilisés en raison de l'effet du filtre temporel déjà présenté. Ces déplacements se sont produits à un pas de temps plus fin que celui du découpage temporel utilisé. Cela porte à quatre le nombre de migrations de Marie non pris en compte. Cependant, si l'on déplace le curseur temporel de dix années, Marie n'est plus considérée comme migrante car, entre 1990 et 2000, elle n'a pas franchi de limite de zone régionale. Ses déplacements régionaux seront vraisemblablement collectés s'ils sont mesurés à l'échelle communale. A l'échelle des communes, Marie a transféré 11 fois sa résidence principale. Enfin, pour la seule région Ile-de-France, Marie aura effectué deux migrations internes entre 1980 et 1990.

L'effet du découpage territorial utilisé pour le passage des données individuelles aux données agrégées de lieu à lieu agit de trois manières sur l'effectif de migrants et sur le nombre de migrations. La première d'entre elles est l'exclusion des migrants qui sont sortis du territoire pendant la période de temps considérée, même s'ils ont entre-temps transféré leur résidence. La seconde est l'exclusion des migrants qui se trouvent, à la date du recensement, dans la même zone que celle de leur résidence antérieure considérée comme le lieu de départ de leur migration : il s'agit là d'une migration interprétée comme interne, alors que ce n'est pas forcément le cas. La troisième porte sur l'agrégation spatiale de données collectées localement – au niveau du logement, notamment – à des échelons territoriaux supérieurs, la commune, par exemple.

1.3.3.2 La constitution de la matrice de flux

Cette procédure d'agrégation consiste à dénombrer le nombre de trajectoires individuelles connues qui franchissent une limite de zone pendant une période, à les fusionner puis à les transférer au niveau supérieur en appliquant une opération mathématique. Pour illustrer ce passage entre les déplacements individuels et les flux agrégés, nous supposons souhaiter disposer de la matrice des migrations de Marie à l'échelle biographique (1960-2040), comme précédemment.

La Figure 1-13, située page suivante, représente la matrice des migrations de Marie réalisée aux échelles régionales et communales de ses transferts de résidence.

On retrouve ici la formalisation générale de la matrice des flux telle qu'elle est présentée dans la section 1.1.1 : sont donc portés en lignes (\rightarrow) l'ensemble des lieux d'origine et en colonne (\downarrow), l'ensemble des lieux de destination. Chaque case de la matrice indique le nombre de migrations mesurées entre un couple de lieux OD. Les cases grisées indiquent des migrations qui se sont produites à une résolution spatiale inférieure à celle qui est présentée dans le tableau : il s'agit de la diagonale de la matrice que nous n'analysons pas, car elle correspond aux flux qui n'ont pas franchi la limite de zone dont il est question.

Sont également renseignées sur la Figure 1-13 : les sommes marginales en lignes (O_i), en colonnes (D_j) telles que présentées sur la Figure 1-3 (page 31), ainsi que le total général des flux élémentaires, naturellement constant aux deux échelles. Sur la matrice des migrations interrégionales présentées sur la partie gauche de cette figure, on observe d'une part, quatre flux *internes* à la région Centre (CEN, CEN). Ces flux *internes* sont aussi qualifiés de flux *intra-zonaux*, *intra-zones* ou *intra-régionaux* car ils n'ont pas franchi de limite de zone, ici celle des régions. D'autre part, un premier flux *externe*, *interzonal*, *inter zone* de la région Centre vers l'Ile-de-France (CEN, IDF) et un second, de l'Ile-de-France vers la région Centre (IDF, CEN) : ces deux flux sont inter régionaux. A l'échelle des communes, ces mêmes flux sont intercommunaux : ils ont franchi la limite de zone communale ; c'est en particulier le cas du flux (To, An) et (Pa, Ro).

Les flux collectés entre le Centre et l'Ile-de-France sont dits *bilatéraux*, à *double-sens* ou encore *aller-retour*, sans pour autant qu'il soit possible de distinguer quel est le flux *aller* du flux *retour*. En revanche, il n'y a pas de flux entre le Centre et la Lorraine : la case correspondante est portée à zéro. A

l'échelle communale, la lecture de la matrice se fait de la même manière qu'à l'échelle régionale. A noter que l'échelle élémentaire de collecte des données étant la commune (d'après le récit), la diagonale des flux intercommunaux est toujours vide.

Figure 1-13. Les matrices (régionales et communales) issues des migrations de Marie

Echelle régionale						Echelle communale																
Destinations (j)						Destinations (j)																
Origines (i)						Origines (i)																
	CEN	IDF	LOR	ALS	Oi		CEN			IDF			LOR	ALS								
CEN	4	1	0	0	5		Go	Ro	To	An	Pa	Pu	Me	St	Oi							
IDF	1	2	1	0	4	Go	0	1	1	0	0	0	0	0	2							
LOR	0	0	0	1	1	Ro	1	0	1	0	0	0	0	0	2							
ALS	0	1	0	0	1	To	0	0	0	1	0	0	0	1								
Dj	5	4	1	1	11	An	0	0	0	0	1	0	0	1								
						Pa	0	1	0	0	0	1	0	2								
						Pu	0	0	0	0	0	0	1	1								
						Me	0	0	0	0	0	0	0	1								
						St	0	0	0	0	1	0	0	0								
						Dj	1	2	2	1	2	1	1	1	11							

Valeur des flux
 0 : absence de déplacements
 4 : nombre de déplacements (flux)

Nom des communes **Nom des régions**
 Go : La Goinfrerie CEN : Centre
 Ro : Romorantin IDF : Ile-de-France
 To : Tours LOR : Lorraine
 An : Antony ALS : Alsace
 Pa : Paris
 Pu : Puteaux
 Me : Metz
 St : Strasbourg

Françoise Bahoken, 2013.

En pratique, la procédure de génération de la matrice des flux opère de la même façon que dans l'exemple que nous avons présenté, à la différence près que ce n'est pas un mais plusieurs événements qui sont agrégés. C'est pourquoi il est dès lors possible de générer une matrice du nombre de *migrants* et une autre du nombre de leurs *migrations*.

Cette procédure qui consiste à passer d'un ensemble de déplacements individuels à un « paquet », c'est-à-dire à un flux agrégé, est parfois considérée à tort comme une réduction de l'information, une généralisation. Il s'agit pourtant d'un changement conceptuel fort qui se répercute au niveau des trois composantes impliquées dans la mesure des flux. Sur le plan thématique, la procédure d'agrégation consiste à changer d'objet d'étude puisqu'elle conduit au passage du déplacement élémentaire à l'agrégat. Ce changement peut être aussi appréhendé comme une transformation du prisme spatial d'observation de l'objet d'étude : les flux sont généralement analysés au regard des zones d'origine et de destination qu'ils mettent en relation à une certaine échelle, c'est-à-dire dans le cadre d'un raisonnement géographique. Sur le plan géographique, l'agrégation consiste, en effet, à un changement d'échelle d'observation d'un même phénomène.

Qu'il s'agisse d'une matrice de flux estimée en termes d'effectifs de migrants ou de nombre de migrations, la référence à l'individu et ses éventuelles caractéristiques attributaires sont perdues dans la procédure d'agrégation, (Fij) étant désormais un agrégat d'effectifs.

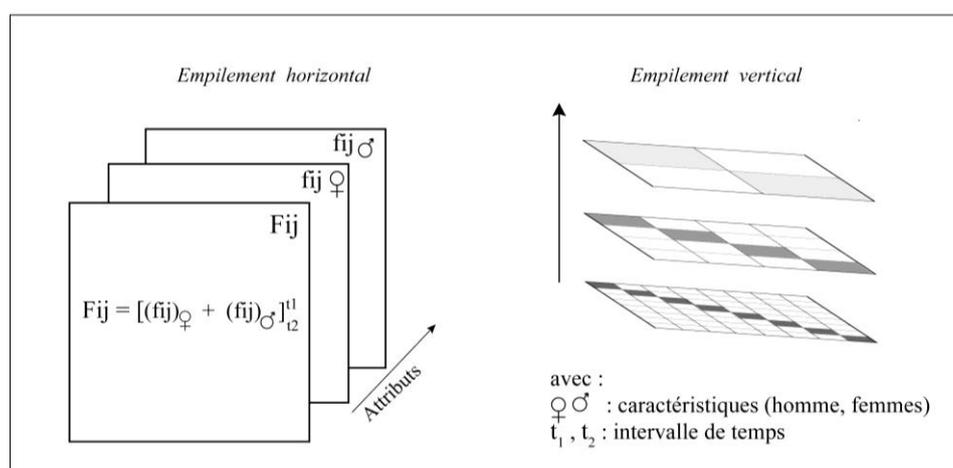
Le fait que la matrice (Fij) soit un agrégat d'effectif indique une possibilité de décomposition des flux concernés en sous-populations statistiques. Cela signifie l'existence potentielle, pour toute matrice

(Fij) portant sur une population totale, de matrices complémentaires qui décrivent une partie de cette population totale par des attributs (le genre, la classe d'âge, ...) qui ont été générés simultanément à la matrice de l'ensemble de la population.

1.3.3.3 La constitution de matrices attributaires complémentaires

Les matrices de flux agrégés (Fij) dont on dispose pour l'analyse résultent en fait d'une agrégation de données déjà disponibles en « paquets » et pour lesquelles les caractéristiques sémantiques des individus ont été perdues dans la procédure d'agrégation. Cependant, ces caractéristiques sont parfois conservées dans des matrices complémentaires (fij) dont l'ensemble correspond au total (à la matrice totale). Trois possibilités sont envisageables : le cas de *sous-matrices* plus/ou moins agrégées dans l'espace et le cas de *sous-matrices* segmentées en catégories sociales ou temporelles, comme illustré sur la Figure 1-14.

Figure 1-14. Les matrices complémentaires



La segmentation sur le plan social correspond, par exemple, à un découpage selon le genre : une première matrice $(fij)_1$ ne porte que sur les femmes et une seconde sur les hommes $(fij)_2$, la matrice d'ensemble étant égale à leur somme $(Fij)=(fij)_1+(fij)_2$. De même la segmentation temporelle (tij) correspond potentiellement aux différentes dates auxquelles les données sont disponibles pour un même espace.

Quelles que soient les caractéristiques de ces matrices complémentaires, toutes permettent l'analyse de la structure des échanges en fonction de catégories (thématiques, temporelles, ...) plus précises que les données d'ensemble : ces catégories correspondent à des groupes sociaux ou économiques, à des produits commerciaux de natures différentes, par exemple. Le cas d'empilement vertical des matrices est celui où les données de flux sont disponibles à plusieurs échelles, différenciant ainsi les lieux d'origine et de destination en fonction de critères *d'appartenance territoriale commune* (à une même « région » qui peut être politico-administrative, économique ou encore linguistique, ...). Nous y reviendrons dans le chapitre 7. Malgré leur intérêt, nous n'analyserons pas le cas de ces matrices complémentaires. La matrice (Fij) dont il sera question par la suite sera toujours unique : elle portera sur l'ensemble d'une population humaine, de biens ou de valeurs considérée et ne sera appréhendée qu'à une seule date (ou sur une période donnée) et à une seule échelle géographique. Ces restrictions particulières sont hors quelques cas particuliers³⁹ que nous expliciterons au moment venu.

³⁹ Un cas de matrice catégorielle est analysé dans le chapitre 5 et un cas d'appartenance territoriale commune des lieux d'origine et de destination à une même région sera envisagé dans le chapitre 7.

Conclusion du chapitre

Les méthodes de collecte des données qui portent sur des déplacements individuels, qu'elles soient directes, indirectes ou alternatives ne sont pas spécifiquement élaborées pour le recueil spécifique des informations qui portent sur des mobilités spatiales. Elles sont par conséquent inadaptées et à la source de problèmes qui ne peuvent pas toujours être compensés. En effet, l'une des difficultés de l'analyse géographique spatialisée concerne la « *sujétion* [des chercheurs] *à l'égard des sources de seconde main* » (Haggett, 1973 : 210) ou, au mieux, comme nous venons de le voir, aux sources incomplètes en raison des « *effets de filtrage* » (Grasland, 1997 : 53) réalisés lors de la mesure. Cette sujétion à la mesure engendre une série de problèmes qui tient davantage à l'inadaptation des données qu'à ce que l'on est en mesure d'en faire.

Les matrices de flux sont des agrégats complexes d'événements sur les déplacements se produisant à des niveaux atomiques plus fins. Les valeurs des flux dont on dispose sont effectivement sous-estimées, par essence : « [...] *l'information relative à la mobilité n'est connue que par le décompte de ces flux entre zones géographiques [...] la mobilité est toujours au plus égale et souvent très inférieure à la mobilité totale [...] elle dépend de la finesse du maillage ou découpage en zones géographiques* » (Pumain, Saint-Julien, 2001 : 14).

On doit donc toujours considérer (d'après C. Grasland)⁴⁰ le hiatus existant entre le concept de flux (Fij) et la réalité statistique de sa mesure et cela, indépendamment du fait que ce dernier n'interfère en rien dans la construction de la carte de flux (d'après C. Zanin).

Par conséquent, il n'est **pas possible d'interpréter une matrice sans reconstituer à la fois le processus de collecte atomique et la méthode d'agrégation employée**. En effet, les problèmes d'ordre méthodologique sont généralement « [...] *plus importants lorsque l'on analyse les flux, beaucoup sont connus et une compensation peut être réalisée. Le coefficient de corrélation (Pearson) des flux interrégionaux qui émanent de deux sources différentes [...]* » (Dorling, 1991 : 105) *peut être utilisé pour estimer la fiabilité des données*. Ces problèmes sont, de manière générale, plus importants lorsqu'il s'agit de flux.

La thématique du flux est, en outre, cruciale, car elle conduit à des procédures très différentes de mesure, de conceptualisation et d'analyse du flux (flux migratoires / flux affectés sur réseau) dont on montrera qu'ils n'interfèrent pas sur la méthode de représentation graphique. C'est pourquoi le second chapitre présente une typologie des modèles de représentation graphique des matrices de flux.

⁴⁰ Cette mention est révélatrice des débats, souvent animés, en réunions de travail, entre mes deux directeurs, selon que le raisonnement était mené au niveau statistique ou cartographique.

2^{ème} CHAPITRE. Les modèles de la représentation graphique des matrices de flux : la matrice ordonnée, le graphe et la carte

« Si ton œil était plus aigu tu verrais tout en mouvement : comme le papier enflammé se déforme, toute chose se défait et se déforme ».

Friedrich Nietzsche (~1889⁴¹).

Introduction du chapitre

Ce second chapitre a pour objectif de poser clairement ce qui est une carte de flux au regard des autres possibilités de représentation graphique des matrices. Pour ce faire, nous présentons les trois modèles graphiques suivants : la matrice ordonnée, le graphe et la carte de flux. Nous utilisons le terme de modèles graphiques car ces diagrammes correspondent, comme l'indiquent P. Haggett (1965, 1973) et F. Durand-Dastès (2001), à la représentation – sous la forme d'une image – de données numériques nécessairement simplifiées. Ils permettent le passage du tableau de données (voir Figure 1-2, page 30) à sa représentation, à l'aide de symboles conventionnels.

Ces modèles sont une formalisation graphique qui donne à voir une structure des relations observées entre des lieux, les motifs (*patterns*) des flux. Ils ont été générés à des fins de description des données par l'image, de recherche d'explication ou de prédiction des relations et, avant cela, d'exploration. La représentation (carto)graphique de données mathématiques est, dans bien des cas, partie prenante d'une chaîne d'analyse commune conduisant à soulever, valider ou infirmer des hypothèses.

Comme nous l'avons mentionné dans l'introduction générale, la carte de flux n'est pas considérée en tant que telle comme un modèle graphique, elle n'est d'ailleurs que rarement mise en œuvre pour représenter des flux. Elle s'inscrit pourtant dans le paradigme de la visualisation des données relationnelles qui conduit à la représentation, sur un support graphique papier ou numérique, de la structure et de la valeur des relations entre différentes entités spatiales.

Nous avançons une hypothèse pour expliquer cette situation secondaire de la carte de flux dans le champ de la visualisation graphique des données statistiques. De notre point de vue, la carte de flux serait considérée comme une variante d'un modèle plus général de représentation des structures relationnelles, ce qui explique son absence – en tant que modèle – dans les classifications et la rareté des innovations dont elle fait apparemment l'objet. La carte de flux n'intéresserait pas le champ de la visualisation des données statistiques pour des raisons qui nous apparaissent, pour l'essentiel, liées à des difficultés d'ordre technique et cognitive relatives à la perception des relations. L'ampleur des contraintes graphiques à résoudre pour sa réalisation n'en fait pas un objet facilement manipulable. Elle ne serait de fait pas attractive, ce qui explique le peu de développements autour de sa figure ainsi

⁴¹ Cette citation, souvent mentionnée partiellement, semble extraite de *Fragments Posthumes*, des carnets de textes rédigés par l'auteur entre 1850 et 1889, mais non publiés.

que les stratégies d'évitement qui la caractérisent, y compris de la part des géographes et des cartographes.

Cette hypothèse s'inscrit, par conséquent, dans **l'idée d'une double difficulté intrinsèque à la réalisation d'une carte de flux** : d'une part, la difficulté de représenter un flux (Fij) par voie cartographique sur un support papier – ou même numérique – et d'autre part, celle de manipuler un tableau dont la structure relationnelle des entités est, par essence, multidimensionnelle (voir chapitre 1). En effet, la représentation graphique de données qui portent sur des relations entre des entités suppose de concevoir – étape d'un processus d'abstraction – puis de représenter, à l'aide de symboles et de principes⁴², les relations et la structure qui existent entre deux objets. Ce processus revient à considérer un géo-dipôle⁴³ ou couple de lieux et la valeur du lien qui les unit, c'est-à-dire trois objets (un lieu d'origine, un lieu de destination et un lien valué) dont les caractéristiques sont archivées sous la forme d'une matrice de flux.

Ce second chapitre a ainsi pour objectif de présenter les trois modèles graphiques de la représentation des matrices, de manière successive, comme autant de possibilités pseudo-équivalentes, souvent complémentaires, de représentation graphique des données de flux. Nous les présentons en adoptant une trame similaire pour chacun d'eux afin de faciliter leur comparaison. Chacune des sections est composée de quatre sous-sections qui portent successivement sur : 1) la présentation du (ou des) objectifs de la figure ; 2) sur les principes de sa construction qui sont suivis 3) de la présentation des principales contraintes et limites graphiques à leur élaboration. Celles-ci sont, en pratique, autant de difficultés qu'il convient de dépasser, autant de défis à relever qui ont conduit, dans bien des cas, 4) à la proposition de variantes au modèle général qui font l'objet de la quatrième sous-section. La carte de flux fait toutefois l'objet d'un développement plus important, puisqu'elle constitue l'objet de cette thèse.

Nous illustrons, dans la mesure du possible, la présentation de ces trois modèles graphiques, ainsi que leurs éventuelles variantes, avec les données disponibles dans une matrice ancienne : celle des flux inter districts de l'ex-Tchécoslovaquie (1989-1990), dont le découpage territorial est présenté sur la Figure 2-1.

L'échelon le plus fin auquel nous ayons accès est celui des 114 districts (*okrej*), qui date de 1961. Il peut être agrégé en douze voire en dix régions administratives (les *kraj*), lorsque Prague est associée à la Bohême-Centre et Bratislava avec la Slovaquie-Ouest. Ces limites régionales suivent à peu près celles des trois anciennes républiques historiques que sont, de l'ouest à l'est : la Bohême qui deviendra la république Tchèque (52 districts), la Moravie-Silésie (24 districts) et la Slovaquie (38 districts). Le troisième niveau est celui des frontières des républiques socialistes Tchèque et Slovaque en vigueur jusqu'en 1989 et le quatrième échelon est celui du périmètre de l'ex-Tchécoslovaquie de 1990. A ces découpages administratifs correspondent « [...] *des limites historiques fossiles qui avaient suivi le même tracé au cours du temps* » (Bopda et al., 1994 : 108). Les limites linguistiques apparaissent toutefois plus perméables que les limites administratives, elles sont, en effet, susceptibles de se modifier « [...] *au cours du temps à la faveur des migrations, des mariages et du renouvellement des générations* » (ibid).

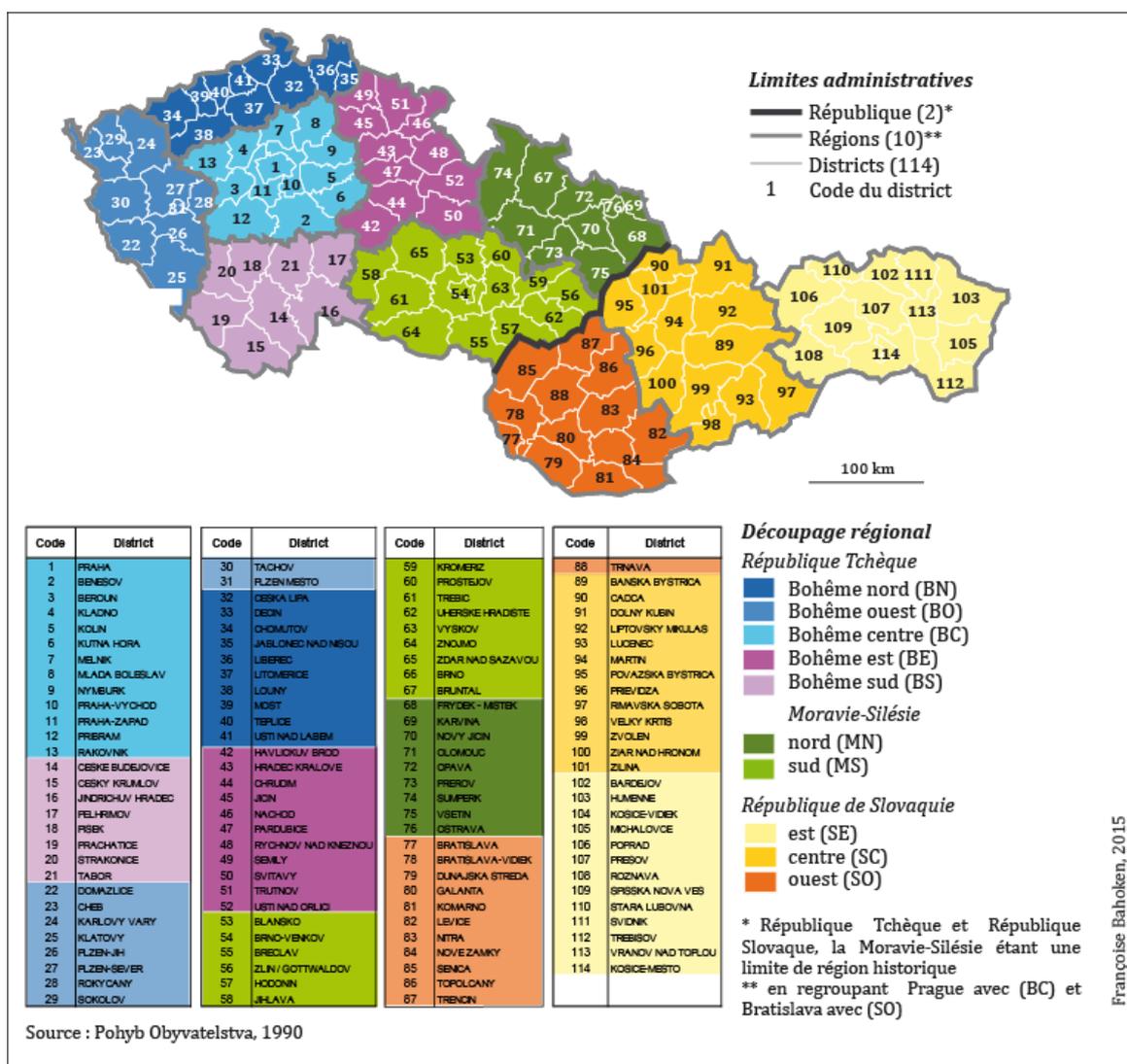
Le choix d'utiliser cette matrice est de nature non seulement méthodologique mais également personnelle. La série de matrices de C. Grasland qui portent sur l'ex-Tchécoslovaquie décrit des

⁴² La seconde section du chapitre 4 porte spécifiquement sur la conception de la carte de flux et la sémiologie cartographique.

⁴³ Une présentation des aspects conceptuels de la représentation des objets géographiques et de leurs relations est disponible dans Goodchild et al. (2007).

données de mobilités résidentielles⁴⁴ qui sont harmonisées et stabilisées. Outre leur intérêt thématique ces matrices peuvent être considérées comme un cas d'école et sur le plan pratique, elles sont suffisamment denses pour pouvoir être intéressantes, tout en restant aisément manipulables⁴⁵. C'est pourquoi elles servent de support pédagogique pour l'enseignement de méthodes d'analyse des interactions spatiales et de cartographie de flux⁴⁶ qui ont été formalisées et développées dans le cadre de nombreux travaux de recherche pouvant être utilisés, le cas échéant, à titre de comparaison. Sur le plan personnel, il s'agit de données que nous avons beaucoup manipulées, pour apprendre et comprendre, avant de faire le choix, pour illustrer les méthodes proposées en seconde partie, de flux plus récents réalisés dans notre département d'origine, l'Isère.

Figure 2-1. Le découpage en 114 districts de l'ex-Tchécoslovaquie de 1990



⁴⁴ Les données issues de l'annuaire démographique Tchécoslovaque (Pohyb Obyvatelstva) ont été collectées et préparées par (Grasland, 1994 : 439).

⁴⁵ C'est généralement une version de ces flux agrégée à l'échelle des dix régions qui est utilisée pour l'enseignement.

⁴⁶ Le support de cours en question (Grasland, Baron, 2006) porte justement sur l'analyse d'une matrice de flux de l'ex-Tchécoslovaquie, en dix régions (pp. 14-19).

2.1 La matrice ordonnée

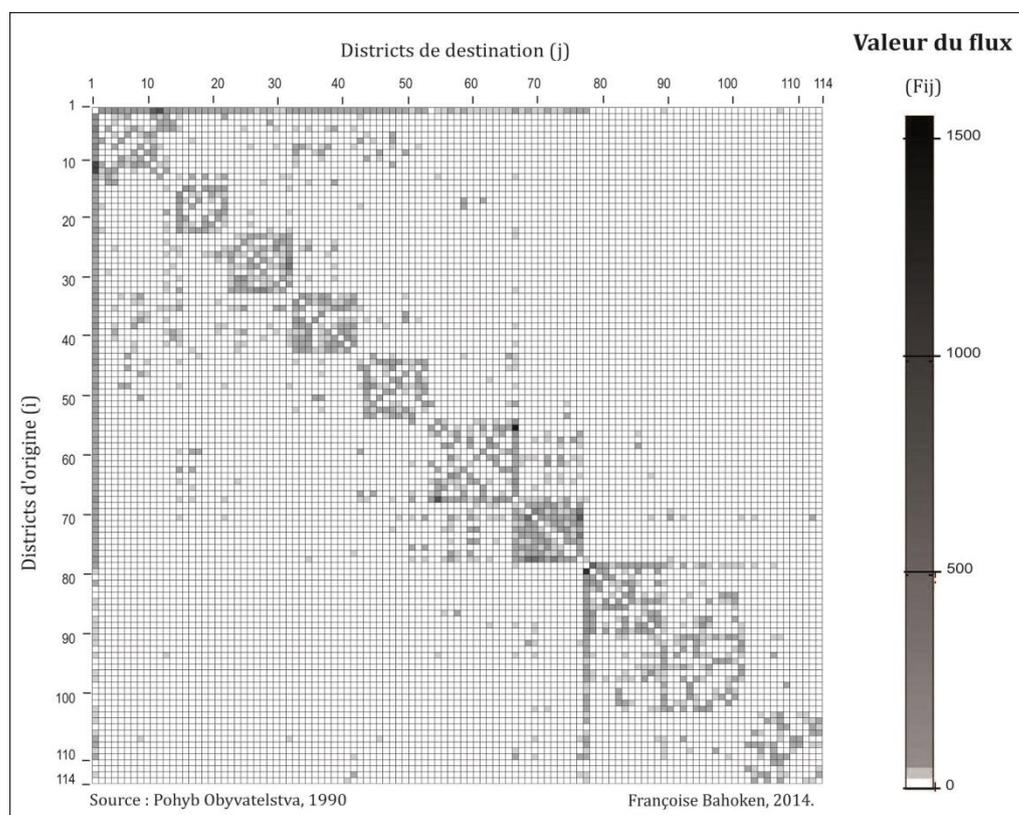
 ⁴⁷ Le mode de représentation en matrice ordonnée – aussi qualifié de matrice bloquée ou de matrice Bertin – a été mis au point d’abord manuellement par J. Bertin, à la fin des années 1960, alors qu’il était chercheur au sein du laboratoire de Graphique de l’Ecole des Hautes Etudes en Sciences Sociales (EHESS). La matrice ordonnée est le résultat de l’application d’une méthode de traitement graphique et visuel de données quantitatives développée à l’heure où les outils informatiques permettant la mise en œuvre de méthodes statistiques de classification n’étaient pas généralisés. Cette méthode reste néanmoins d’actualité, car il est désormais possible de la mettre en œuvre rapidement et plutôt facilement, dans le cadre d’une première visualisation voire d’une exploration préliminaire de la matrice de flux.

2.1.1 L’objectif

L’objectif général de la représentation d’une matrice de flux (F_{ij}) sous la forme d’une matrice ordonnée tient au recodage des valeurs de flux, puis à leur transformation de manière à aboutir à une typologie visuelle qui est ensuite, éventuellement, spatialisée. Le résultat obtenu prend la forme d’une grille composée de cellules de tailles régulières et colorées. Chaque cellule, qui correspond au croisement d’une ligne (i) et d’une colonne (j), est assortie d’une nuance de teinte, selon une gamme chromatique. Cette (nuance de) teinte correspond à une valeur de flux, généralement un stock.

La Figure 2-2 présente la matrice des flux de l’ex-Tchécoslovaquie, ordonnée selon l’identifiant du district (de 1 à 114), en niveaux de gris, une version interactive de cette figure est disponible dans Bahoken (2015).

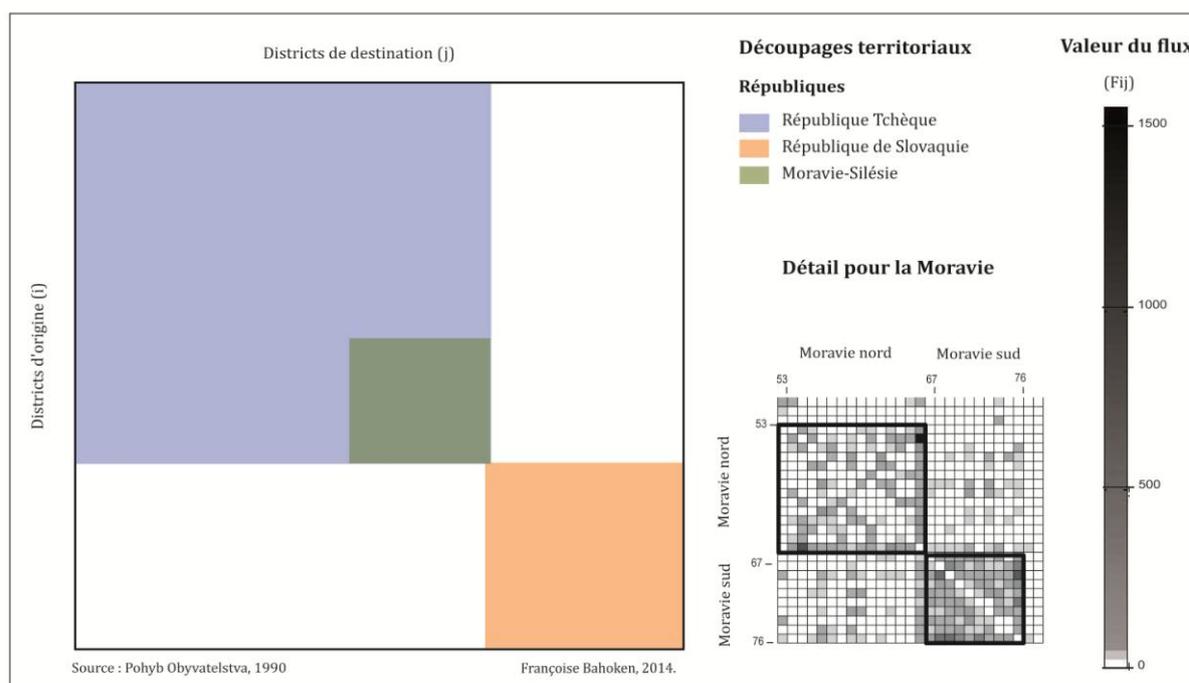
Figure 2-2. La matrice ordonnée de l’ex-Tchécoslovaquie (1990)



⁴⁷ Nous ponctuons le texte d’icônes qui illustrent les modèles graphiques dont il est question dans la section.

On observe une forte concentration des flux à l'intérieur de chaque région (*okrej*) correspondant au chef-lieu. Cette matrice ordonnée met en évidence la structure spatiale régionale de ces flux, soulignant le rôle joué par les frontières correspondantes (limites de zones) et l'antériorité de leurs relations qui se déploient au-delà du découpage administratif de 1991, comme illustré sur la Figure 2-3.

Figure 2-3. Le découpage régional sur la matrice ordonnée de l'ex-Tchécoslovaquie (1990)



Les flux se produisant au sein des deux régions de Moravie nord (MN) et sud (MS), mettent effectivement en évidence la permanence des relations intra régionales (MN, MS) mais aussi inter régionales, dans la limite du périmètre de l'ancienne partie Moravie-Silésie. On observe également la porosité de cette limite historique, visible par l'intensité des cellules correspondant aux flux qu'on observe avec les districts « frontaliers », qui appartiennent à la République Tchèque de 1990.

Le recodage des valeurs de flux en nuances colorées rend la matrice de flux lisible. Elle correspond à un usage particulière de la variable visuelle⁴⁸ *Valeur*⁴⁹, puisqu'elle est appliquée ici à des stocks (et non à une variable de taux ou de rapport). Cela ne s'inscrit pas en faux par rapport aux principes de sémiologie graphique, définis par J. Bertin (1969, 1973) car l'objectif de l'ordonnement de la matrice est de rechercher un ordre dans ses valeurs, « [...] *des relations internes, c'est-à-dire des relations contenues dans les données inventoriées* [...] » (Bertin, 1969 : 70), de manière exploratoire. La posture J. Bertin est alors celle d'un explorateur de données⁵⁰ qui agit à la manière du statisticien

⁴⁸ J. Bertin (1967) a identifié six variables visuelles qui sont : la couleur ; la forme ; la texture-structure (et/ou le) grain ; l'orientation, la valeur et la taille. A noter que la texture-structure peut être appréhendée de deux manières : dans son acception anglaise, comme un synonyme du grain ou bien, à l'inverse, dans son acception francophone dont la construction est différente : la texture correspondant dans ce cas à un mode d'application de la variable visuelle valeur, sous la forme de trame et non d'aplats ; le grain étant la résolution de l'image.

⁴⁹ Les variables visuelles seront présentées en italique, avec une majuscule, quelle que soit leur position dans le texte.

⁵⁰ Un *data scientist* comme on dit, aujourd'hui.

cherchant à classer ses données et non du cartographe cherchant à les représenter efficacement. Aussi élude-t-il les idées de quantité et de proportionnalité entre ses valeurs qui sont, en l'état, illisibles.

Le recodage des valeurs permet de faire émerger puis de visualiser un classement ordonné des valeurs dans le but suivant : « *Montrer que les groupes dégagés forment la moins discutable des réductions possibles de l'information. C'est la discussion des groupements expérimentaux successifs [...] c'est-à-dire l'existence d'une relation interne⁵¹ a priori* » (Bertin, 1969 : 74). On voit bien sur la Figure 2-2 le résultat de l'ordonnement des cellules qui permet de distinguer des « paquets », d'où la dénomination de la figure en matrice ordonnée.

Si la matrice ordonnée a d'abord été développée pour faciliter la réalisation manuelle de typologies spatiales dont le résultat s'apparente à ceux qui sont issus de méthodes statistiques plus classiques de classification (Bertin, 1969), nous pensons qu'elle constitue également un bon moyen de représenter graphiquement les données disponibles dans une matrice de flux et cela, outre les possibilités exploratoires qu'elle autorise par ailleurs. C'est pourquoi nous l'élevons au rang de modèle élémentaire de représentation graphique des matrices de flux.

C'est J.-C. Müller (1991) qui semble être le premier à avoir suggéré l'utilisation de la matrice ordonnée pour représenter des flux, en complément de méthodes graphiques et mathématiques telles qu'elles sont mises en œuvre par J. Bertin (1969) lui-même, ou encore dans les travaux de W. Tobler, sur lesquels nous reviendrons dans le chapitre 3. Sur le plan graphique, la matrice ordonnée est, en pratique, le produit d'une transformation des valeurs de flux en nuances de teintes.

2.1.2 Le principe

Le principe de la matrice ordonnée consiste à recoder visuellement, à l'aide de la *Valeur*, éventuellement assortie à la *Couleur*, les différentes valeurs (i, j) de la matrice. La couleur correspond à une teinte exprimée en niveaux de gris ou dans une gamme de couleurs, une variation de couleurs dans un espace colorimétrique tel que le Rouge, Vert, Bleu (RVB). L'auteur justifie le recours au recodage coloré par la difficulté de distinguer aisément les différentes valeurs numériques d'un tableau croisé, ce qui est préjudiciable à l'analyse : la matrice de données numériques n'autorise pas *a priori* une vision globale du phénomène représenté, à l'inverse des variations de teintes.

Lorsque les cases (i, j) de la matrice correspondent à des données quantitatives⁵², J. Bertin indique ensuite recourir notamment à des barres verticales dont la hauteur est proportionnelle à la quantité concernée. Il utilise enfin un principe de ré ordonnancement des cellules qui consiste à permuter⁵³ visuellement les (m) lignes et les (n) colonnes (j=1, ... n), afin d'aboutir à une typologie visuelle.

L'opération d'ordonnement, conduit à un tri des données de manière à faire émerger les traits caractéristiques de la structure relationnelle des entités. Si elle est nécessaire à une visualisation efficace sous la forme matricielle, elle constitue la principale contrainte à la mise en œuvre de ce modèle.

2.1.3 La contrainte

D'après M. Beguin et D. Pumain (1994, 2003), la visualisation directe du tableau numérique sous la forme d'une image colorée présente l'avantage d'être « [...] *plus facile à mémoriser si un certain*

⁵¹ Ce terme est porté en italique dans le texte de l'auteur.

⁵² Les valeurs quantitatives s'opposent aux valeurs qualitatives nominales et ordinales car elles sont mesurables. Les opérations mathématiques qui leur sont appliquées ont, par conséquent, un sens.

⁵³ Lorsque la permutation est réalisée sur des données qualitatives, l'opération de permutation est qualifiée de diagonalisation. Nous reviendrons sur ce terme dans le chapitre 3, car il convient de ne pas le confondre avec son acception mathématique.

ordre existe déjà entre les lignes et les colonnes de celui-ci » (Beguin, Pumain, 1994 : 163). Il est donc impératif de réordonner les valeurs, ce qui est une contrainte compte tenu de la difficulté de l'exercice. J. Bertin (1973) a expliqué la mise en œuvre de la procédure de ré ordonnancement qu'il propose par l'utilisation d'un matériel particulier nommé Jeu de dominos (Bertin, 1973 : 169 ; Blin, Bord, 1993 : 189). Le jeu de dominos est utilisé lorsque l'opération de permutation ne peut être réalisée manuellement, n'étant pas facile à mettre en œuvre. Il consiste à assembler des perles de dés colorées qui correspondent aux cases (i, j) de la matrice, sur des pics, qui correspondent aux lignes, puis à les permuter visuellement en observant le codage visuel caractéristique d'une structure qui émerge. Aujourd'hui, peu d'outils permettent une réalisation aisée d'une matrice ordonnée. On peut toutefois citer : le logiciel de calcul matriciel Matlab© ; le module Amado© du logiciel d'analyse statistique Spad© ; l'outil Matrice-Bertin du logiciel de cartographie statistique Cartes&Données© ; les applications d'exploration interactives des matrices discrètes MatrixExplorer© (Henry, Fekete, 2008) et Bertifer © (Perin et al., 2014) de l'équipe Aviz de l'Inria ; des matrices continues issues des travaux de G. et N. Andrienko (2007, 2008) et Wood et al. (2010).

Dans l'exemple de l'ex-Tchécoslovaquie (1990), la matrice est triée selon les codes administratifs par région historique, puis par région administrative. La structuration de cette codification permet de bien visualiser des structures de référence pour les échanges, intra-républiques ou intra-régionales.

Le développement d'applications informatisées d'exploration et de visualisation graphique des matrices a récemment contribué au renouvellement des variantes, la plus courante ayant longtemps été la matrice en points proportionnels.

2.1.4 Les variantes

Les variantes au modèle de la matrice-ordonnée peuvent être regroupées en quatre catégories correspondant respectivement à la matrice en points proportionnels, à la matrice en étoile, aux matrices hybrides et aux matrices de surfaces continues. Elles sont présentées successivement dans les sous-sections suivantes.

2.1.4.1 La matrice en points proportionnels



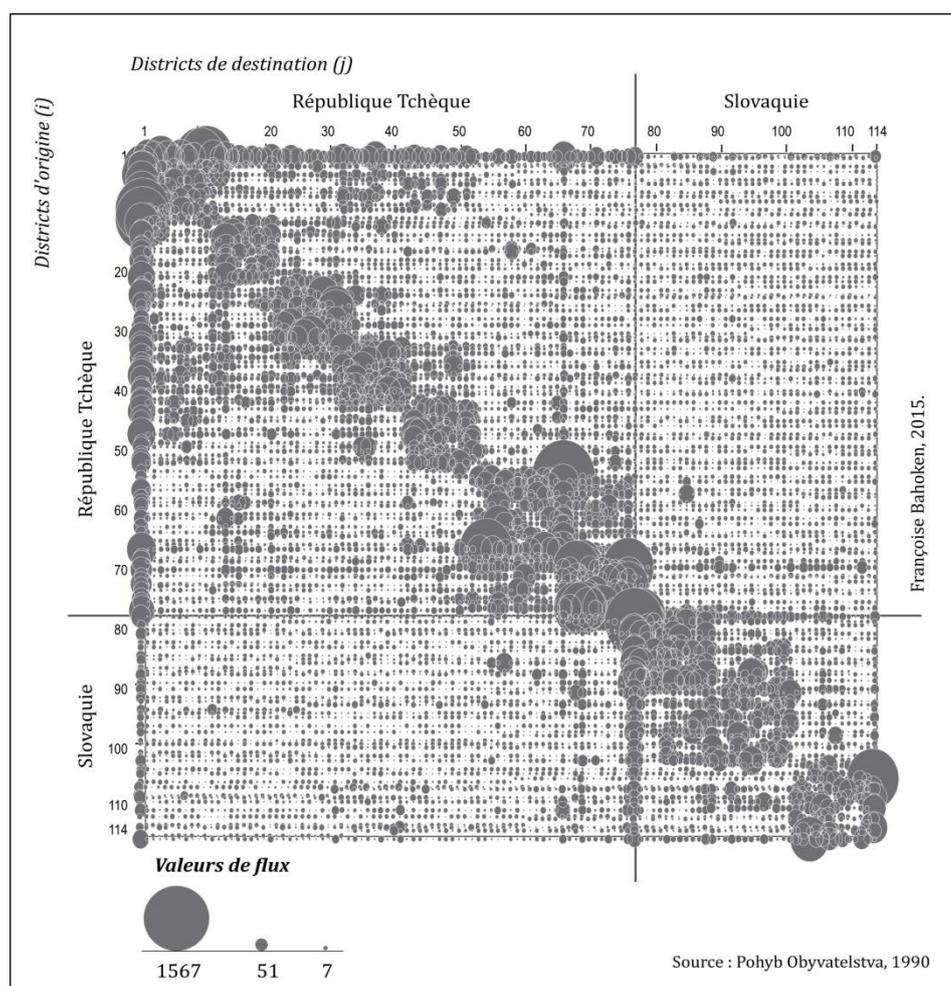
La matrice en points proportionnels, aussi nommée matrice en semis de points, consiste, d'après M. Beguin et D. Pumain (1994), en une représentation directe du tableau numérique sous la forme d'une matrice de cellules régulières formée de points dont la surface est proportionnelle⁵⁴ à la valeur du flux qu'ils représentent (voir Figure 2-4).

L'intérêt d'une représentation en symboles proportionnels tient à l'amélioration qualitative apportée par la visualisation des différences de proportionnalité des valeurs, elles illustrent des inégalités de poids des valeurs brutes qui portent sur des effectifs. Cet avantage peut toutefois se transformer en inconvénient si la représentation est réalisée dans le cadre d'une analyse géographique, en raison de *l'effet de taille*⁵⁵ du maillage qui n'est pas pris en compte dans la représentation sous la forme de matrice puisque les cellules présentent la même taille. Ainsi, « *Si l'objectif de la représentation est de visualiser ces volumes, l'image est utile, sinon il convient de procéder à un traitement préalable de l'information contenue dans le tableau* » (Beguin, Pumain, 1994 : 163), nous y reviendrons dans le chapitre 3.

⁵⁴ Pour cela, il faut utiliser $\sqrt{(Fij)}$ et non (Fij) pour calculer la valeur du rayon du cercle ou indiquer cette valeur dans la matrice avant sa représentation graphique, sous Rstudio (c), par exemple.

⁵⁵ *L'effet de maillage* se manifeste dès lors que les unités spatiales d'une zone d'étude ont une taille et une forme hétérogène, ce qui est généralement le cas dans la réalité. Nous présenterons les aspects graphiques de cet effet dans la section 3 de ce chapitre.

Figure 2-4. La matrice en points proportionnels de l'ex-Tchécoslovaquie (1990)



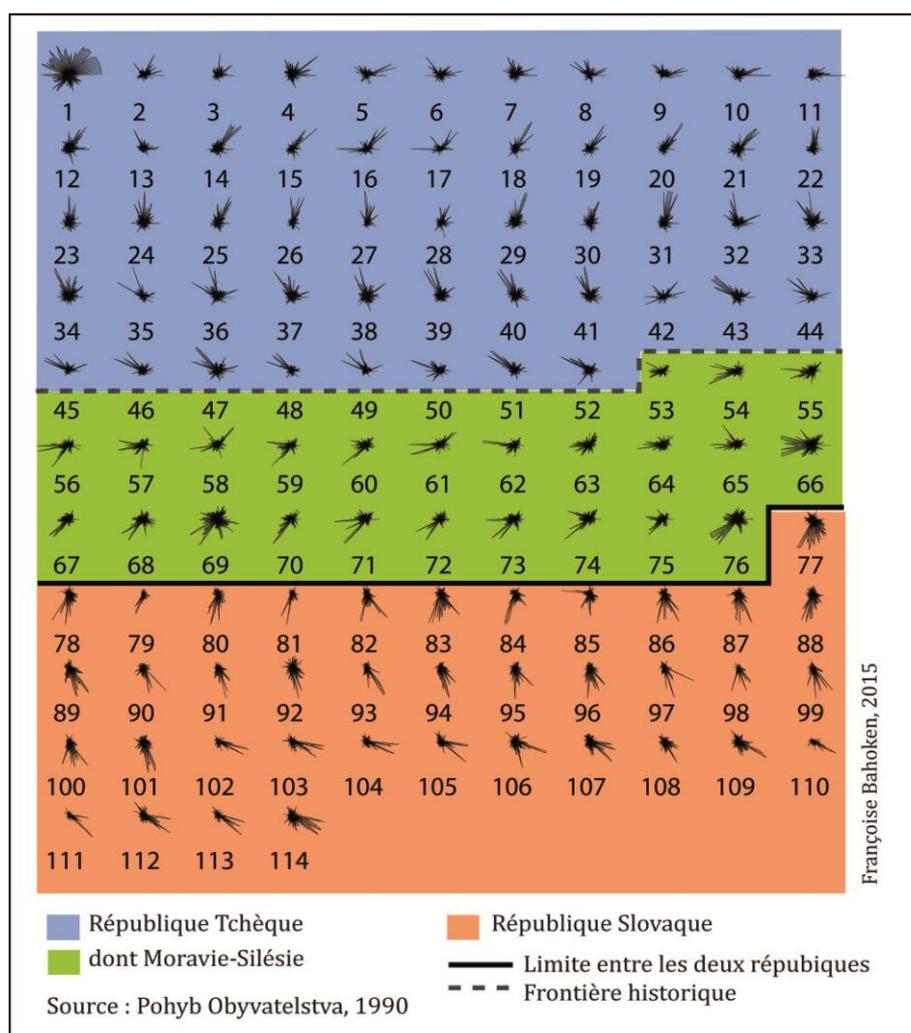
2.1.4.2 La matrice en étoile

Une seconde variante de la représentation en points proportionnels est bien connue des informaticiens. Réalisée du point de vue des lieux, elle consiste à représenter le système de flux de chacun d'eux sous la forme d'étoiles (Ward, 2002).

Ainsi, à chaque ligne (ou lieu d'origine) de la matrice est associée une étoile, dont le nombre de branches correspond au nombre de lieux de destination des flux, la longueur des branches dépend de la valeur du flux vers le lieu de destination correspondant. Ainsi, les étoiles les plus fournies correspondent aux lieux qui émettent vers de nombreux lieux, celles qui sont les plus amples émettent les flux les plus forts, décrivant ainsi une morphologie des lieux variée.

La Figure 2-5 représente la matrice en étoiles des flux de l'ex-Tchécoslovaquie (1990). Cette figure, aussi séduisante qu'elle puisse apparaître, pose un problème lié à l'introduction de l'orientation du flux, de sa direction. En effet, sur la Figure 2-5, les liens ne semblent pas orientés en fonction d'une référence géographique ou cardinale : on ne sait donc pas quelle est la signification des angles. Après réflexion, il semblerait que la position des lieux de destination sur un cercle environnant le lieu d'origine correspondrait à leur ordre dans la liste des identifiants des districts (de 1 à 114) et non à une référence angulaire sur le cercle. Cela pose question quant à l'applicabilité et à l'utilité des procédés graphiques générés en dehors de leur contexte thématique d'application. La question n'est effectivement pas tant de produire une belle figure que de réaliser une image censée.

Figure 2-5. La matrice en étoiles de l'ex-Tchécoslovaquie (1990)



La principale difficulté de la mise en œuvre des matrices en points proportionnels (ou en étoiles, d'ailleurs) consiste en l'ajustement de la taille du plus petit et du plus grand symbole, étant entendu que le premier doit rester lisible et le second restreint, autant que possible, à la dimension de la cellule. C'est pourquoi d'autres solutions de représentation matricielle ont pu être proposées, à la place des symboles proportionnels. Wood et al. (2011) suggèrent, par exemple de réaliser une matrice ordonnée qu'ils qualifient d'« *OD-map* »⁵⁶ quand Slingsby et al. (2009, 2010) en proposent une sous la forme d'arbres hiérarchiques (Treemaps).

Ces propositions, toutes issues du champ de l'*InfoViz*, sont en fait des variantes au modèle général de la matrice ordonnée. Leur intérêt est davantage technique qu'analytique, il réside dans l'algorithme permettant leur réalisation, et dans leur esthétique. Elles sont d'ailleurs souvent présentées dans le cadre d'applications⁵⁷ informatiques permettant d'explorer les matrices et de les visualiser de différentes manières, dans l'objectif de remédier aux procédés plus classiques de représentation (par exemple, sous la forme de diagrammes liens-nœuds qui fait l'objet de la section suivante). Cependant, pour autant qu'elles soient esthétiques, ces solutions aboutissent hélas à des résultats dépourvus de

⁵⁶ Une présentation vidéo de la méthode de construction de l'*ODmap* est disponible sur Vimeo, à l'adresse suivante : <http://vimeo.com/45078794> [lien vérifié le 02/08/2015].

⁵⁷ L'application Hide (*Hierarchical Visualization Expression Language*) est en un bon exemple, elle est disponible gratuitement à l'adresse suivante : <http://gicentre.org/hide/> [lien vérifié le 02/08/2015].

sens, sinon assez limités. L'« *OD map* » de Wood et al. (2011) ne permet pas une analyse du cœur de la matrice (des liens ou des flux), mais seulement d'une partie de l'effet des flux sur les lieux. Les auteurs ont d'ailleurs précisé les points suivants : d'une part, que pour un comté irlandais donné (qui correspond au découpage territorial utilisé dans leur analyse), il n'est pas possible de savoir d'où proviennent les flux qui y arrivent ; d'autre part, que leur *ODmap* ne permet pas d'analyser la structure des échanges en termes de voisinage, mais plutôt en termes d'évolution, lorsque les données sont disponibles.

2.1.4.3 La matrice hybride

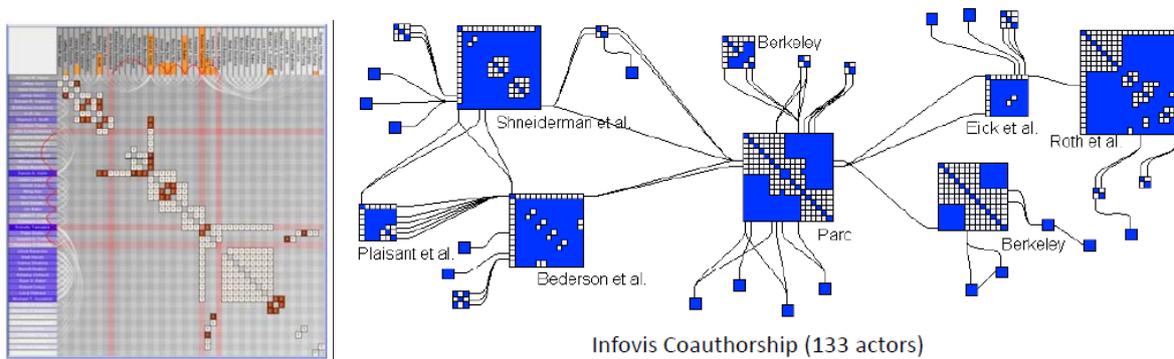
Le procédé de la matrice ordonnable, tel qu'il est mis au point par J. Bertin vers 1969 sur des données numériques matricielles, a été transféré ces dernières années au champ de la visualisation des réseaux sociaux (Ghoniem et al., 2004 ; Henry, Fekete, 2009 ; Guo, 2009 ; Wilkinson, Friendly, 2009 ; Wood, Dykes, 2010 ; Wood et al., 2011).

Pour ces chercheurs en Informatique graphique, le modèle graphique matriciel est un objet de recherche à part entière qui prend différentes acceptions. N. Henry et J.-D. Fekete (2008) parlent de « *matrices* » ou de « *matrice-Bertin* » ; D. Guo (2008) utilise plutôt l'expression « *matrice OD* » pour matrice Origine-Destination et Wood et al. celle de « *ODmap* » déjà mentionnée. On rappelle que J. Bertin (1967) utilise le terme de matrice ordonnable.

C'est ainsi que le potentiel exploratoire du modèle de la matrice ordonnée a été démontré dans le cadre de la recherche de structures internes des grandes matrices d'adjacence, celle des communautés par exemple, (Fekete, Henry, 2008). Certains auteurs l'ont utilisé conjointement à des traitements statistiques d'agrégation (Guo, 2007, 2009) ; d'autres ont combiné le modèle matriciel avec le modèle « lien-nœuds » (voir section 2.2) pour maximiser les possibilités d'exploration de la matrice (Ghoniem et al, 2004). Les versions proposées par (Henry, Fekete, 2001, 2006, 2009 ; Henry et al., 2007) s'appuient sur différents algorithmes de ré ordonnancement optimisés pour des *matrices d'adjacence* et non pour des *matrices de flux*.

La Figure 2-6 illustre ce mode de représentation appliqué à une matrice de connectivité décrivant un réseau d'acteurs publiant ensemble. Il est aisé d'imaginer le même type de figure pour des matrices de flux.

Figure 2-6. Exemples de matrices ordonnées hybrides



Sources : Henri, Fekete, 2006, 2008.

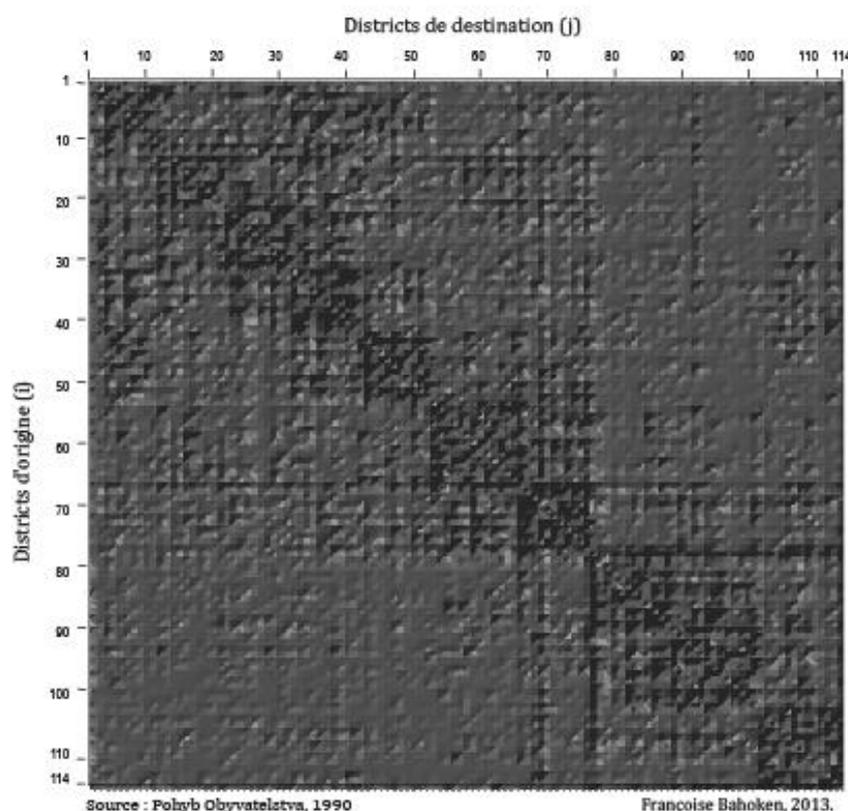
Ces matrices hybrides ont surtout une visée exploratoire à plusieurs niveaux d'analyse qui est renforcée par l'interactivité des figures. En effet, les figures permettent l'exploration des données à deux niveaux structurels : la forme matricielle apparaît réservée à la vision d'ensemble (ou globale) et

la forme relationnelle en liens-nœuds permet une vision locale des acteurs et du sous-réseau de relations dans lequel ils sont impliqués.

2.1.4.4 La matrice en volume

Le quatrième type de variante au modèle de la matrice ordonnée prend une forme en volume et non plus discrète, comme les précédentes. Ce mode de représentation fait suite à la proposition (Tobler, 1987 : 156) de recourir à une approche volumique dès lors que la matrice est dense et que l'on ne souhaite pas opérer de sélection ou de réduction des données *a priori*. La matrice peut ainsi être représentée comme un relief où chaque valeur correspond à une altitude. La Figure 2-7 illustre ce procédé appliqué à l'ex-Tchécoslovaquie.

Figure 2-7. La matrice en 3D de l'ex-Tchécoslovaquie (1990)



Ce mode de représentation volumique (il aurait aussi pu être prismatique) d'une matrice de flux émerge suite à l'impossibilité de visualiser clairement les *matrices denses*, lorsque l'on ne souhaite pas opérer de sélection *a priori* des données. La visualisation des masses de données relationnelles serait par conséquent soumise à un impératif de sélection au préalable de leur représentation, une hypothèse qui mérite d'être vérifiée.

Cette proposition s'inscrit dans un changement de paradigme de la représentation des données disponibles dans une matrice : le passage en trois dimensions. Il témoigne des limites de l'approche graphique discrète, caractérisée par la difficulté de visualiser l'exhaustivité des relations entre les entités d'une matrice dense ; à la nécessité d'une sélection des données ainsi qu'à la nécessité d'une réduction préalable des données.

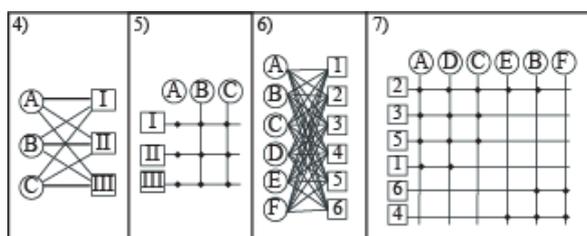
La matrice ordonnée est efficace pour explorer les données et faire émerger une structure spatiale, si le tri préalablement réalisé est pertinent. Elle ne permet pas de localiser les différents espaces, ce qui

pose un problème spécifique à l'analyse géographique. Ce modèle n'est toutefois pas dénué d'intérêt car il répond bien à l'objectif de la figure réalisée, d'illustrer des données relationnelles.

Le regain d'intérêt qu'on accorde aujourd'hui à la représentation matricielle est qu'elle est appréhendée comme une « [...] *alternative intéressante à la métaphore des nœuds et des liens* » (Ghoniem et al, 2004), d'où le développement des variantes hybrides que nous avons déjà évoquées ; celles là sont d'ailleurs souvent réalisées en complément d'un modèle graphique plus populaire à savoir : le diagramme lien-nœud.

D'après J. Bertin (2000), la matrice permet de détecter les structures et les « *relations aberrantes* » des matrices denses, qui présentent un nombre important de liens ; elle peut être représentée par des lignes et des points (voir Encart 2-1).

Encart 2-1. J. Bertin (2000) face à la représentation des matrices



« Une donnée est une relation entre deux éléments. En face, le plan nous offre des points et des lignes. On peut donc représenter les éléments par des points et les relations par des lignes (4). On construit un RÉSEAU. Les dimensions X et Y de l'image ne sont pas significatives. On peut aussi représenter les éléments par des lignes et les relations par des points (5). On construit une MATRICE. Les dimensions X et Y ont chacune une signification. [...] Si le RÉSEAU transcrit parfaitement l'ordre topographique, il est très limité devant les tableaux ordonnables : peut-on par exemple découvrir la relation aberrante en (6)? Elle apparaît immédiatement dans la matrice (7). La MATRICE construit le support inconscient de la réflexion, souligné par l'universalité du 'tableau à double entrée' et des 'procédures de reclassement' ».

Source : J. Bertin, 2000.

Note : ce que J. Bertin nomme « *réseau* », nous le qualifions de « *graphe* » qui correspond au terme qu'il convient d'employer pour évoquer la représentation graphique d'un réseau.

A noter que la matrice dont il est question ci-dessus n'est pas *fermée* : les individus portés en lignes diffèrent de ceux qui sont en colonne : le graphe est bipartite ou *dual*.

2.2 Le diagramme liens-nœuds ou graphe



Le diagramme liens-nœuds, plus connu sous le terme de *graphe*, est le mode de représentation graphique généralement privilégié pour visualiser des données relationnelles. Le graphe est effectivement réputé parfaitement adapté à la visualisation des données relationnelles, grâce à sa formalisation graphique, simple et efficace, sous la forme de lignes et de points.

Le terme de graphe revêtant plusieurs acceptions⁵⁸, celle dont il est question ici est relative à un réseau de relations. Il caractérise alors un mode de représentation des matrices parmi d'autres, (le tableau croisé, la liste de sommets, la matrice ordonnée). Si l'acceptation commune du terme se réfère à la

⁵⁸ Nous proposons une présentation du terme de graphe, dans le billet *G comme Graphe* du carnet de recherches du groupe fmr, disponible à l'adresse suivante : <http://groupefmr.hypotheses.org/2801> [lien vérifié le 02/08/2015]

théorie des graphes, le *graphe* désigne également une fonction mathématique. Dans la théorie des ensembles, on parle de graphe ensembliste pour désigner une structure formée de deux ensembles : celui des sommets et celui des liens, qu'ils soient munis ou non d'attributs particuliers. Ce graphe ensembliste désigne donc l'ensemble des couples d'entités qui vérifient une relation donnée, ainsi que la représentation de cette relation sous la forme d'un schéma, le *graphe*, composé de lignes et de points.

Tout graphe (G) se « [...] définit par son ordre qui correspond au nombre de sommets et par sa taille, définie par le nombre de liens [...]. Dans un graphe non planaire G de taille N, le nombre de liens K est au minimum de 0 et au maximum $N(N-1)/2$. Un graphe $G_{N,K}$ est dit complet (complete) si $K = \frac{N}{2} = N(N-1)/2$, il est noté K_N » (Beauguitte, 2010 : 1-2).

2.2.1 L'objectif

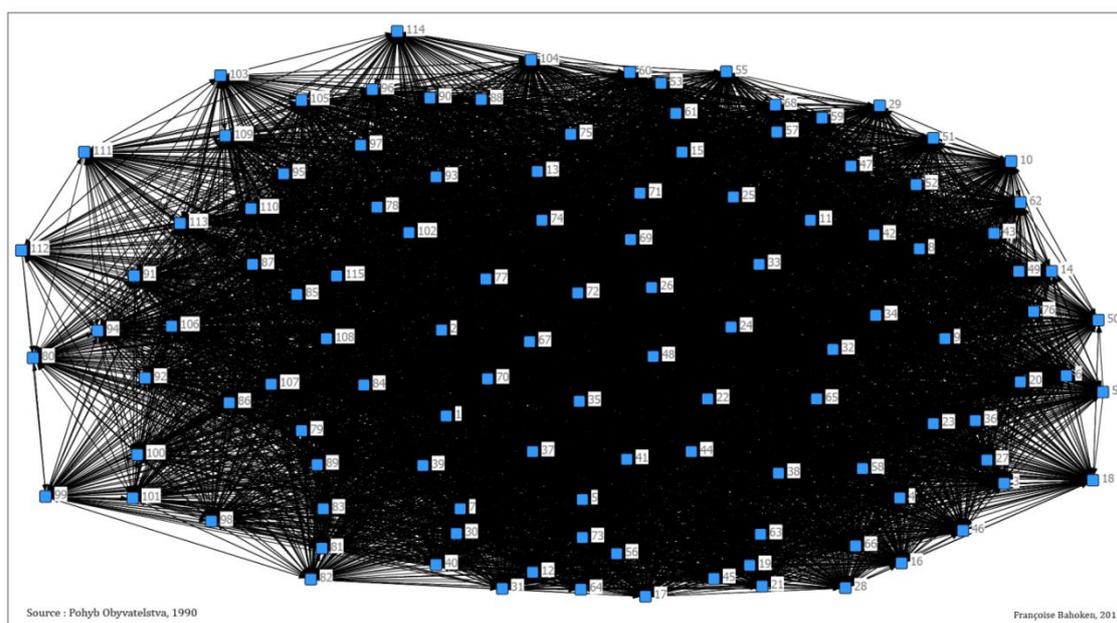
L'objectif premier de la représentation d'une matrice de flux sous la forme d'un graphe est de visualiser la structure des relations qui existent entre des entités. Elle transforme pour cela les valeurs (F_{ij}) de la matrice en objets graphiques : des points et des lignes qui correspondent respectivement aux objets analysés (des acteurs ou des lieux, par exemple) et à la représentation de leur mise en relation.

Le graphe étant une figure de la théorie des graphes, il est d'abord utilisé pour visualiser les matrices d'adjacence (L_{ij}), avant de servir également pour les matrices de flux (F_{ij}). Le graphe est une structure graphique formée de N sommets, qui correspondent aux entités et de L arcs qui traduisent leurs liens. Les sommets du graphe sont, en fonction des cas, l'origine (i) d'un arc (L) ou bien sa destination (j). Chaque arc (L) est caractérisé d'une part, par un couple (i,j), dont le lien est parfois orienté : de i vers j L(i,j) ou à l'inverse, de j vers i L(j,i) et d'autre part, par une mesure de *longueur* qui apparaît visuellement dans la représentation (voir Figure 2-8) mais qui n'est pas définie au départ. Les sommets peuvent être, alternativement, l'origine d'un premier arc et la destination d'un second arc : dans ce cas, ces arcs sont des arêtes et les sommets correspondant sont dits voisins ou adjacents.

La Figure 2-8 présente le graphe de l'ensemble des relations disponibles dans la matrice de flux de l'ex-Tchécoslovaquie (1990). Cette matrice étant valuée (F_{ij}), ses valeurs ont fait l'objet d'une transformation binaire (F_{ij}=1) ou (F_{ij}=0) afin que ne puisse être représentés que l'existence d'un lien entre deux districts (qui correspondent donc aux cellules présentant la modalité 1 de la matrice). Les sommets sont représentés par des carrés bleus qui sont affectés d'un label correspondant à l'identifiant du district présenté sur la Figure 2-1 (page 57).

La Figure 2-8 décrit les 10 356 liaisons (hors diagonale) correspondant aux valeurs pouvant être représentées graphiquement. Il s'agit de l'information réellement disponible pour l'analyse parmi le nombre k de liaisons théoriquement possible (égal ici à 12 996). Les cellules de la matrice non concernées sont portées à zéro, car elles caractérisent une absence de relation entre les couples de lieux correspondants.

Figure 2-8. Le graphe des relations observées de l'ex-Tchécoslovaquie (1990)



Le graphe de la Figure 2-8 prend la forme d'un amas de liens enchevêtrés et superposés qui caractérise un problème de *surcharge graphique*. Ces liens qui sont plus ou moins longs, sont ponctués par des sommets que l'on ne fait que deviner. Cette surcharge correspond à *l'effet spaghetti* qui se manifeste, sur un graphe, au niveau des liens et des lieux : les premiers se croisent et les seconds se chevauchent (Henry, Fekete, 2008 : 62). Sur la Figure 2-8, 10 356 liens entrecroisés sont représentés. Le graphe d'une matrice dense étant, de manière générale, difficile à interpréter, un second objectif précède généralement à sa construction à savoir : la réalisation d'une figure lisible, qui mette en évidence l'une ou l'autre structure de la matrice représentée.

Plusieurs approches peuvent être mises œuvre pour atteindre ce second objectif : la première est fondée sur la détection de *communautés* ou de *régions*, qui correspond à la réalisation d'une typologie des acteurs en classes d'appartenance. Une *communauté* étant effectivement formée par un ensemble d'individus qui sont caractérisés par un (voire plusieurs) attribut(s) commun(s). Il s'agit donc d'individus qui partagent une caractéristique commune : les acteurs qui appartiennent à une même communauté sont plus liés, plus proches et même plus similaires entre eux qu'avec ceux des autres *communautés*. Ils définissent donc un groupe, ou une classe, qui peut être analysé en tant que tel, moyennant un changement conceptuel de l'objet d'étude qui tient au passage de l'individu à l'agrégat.

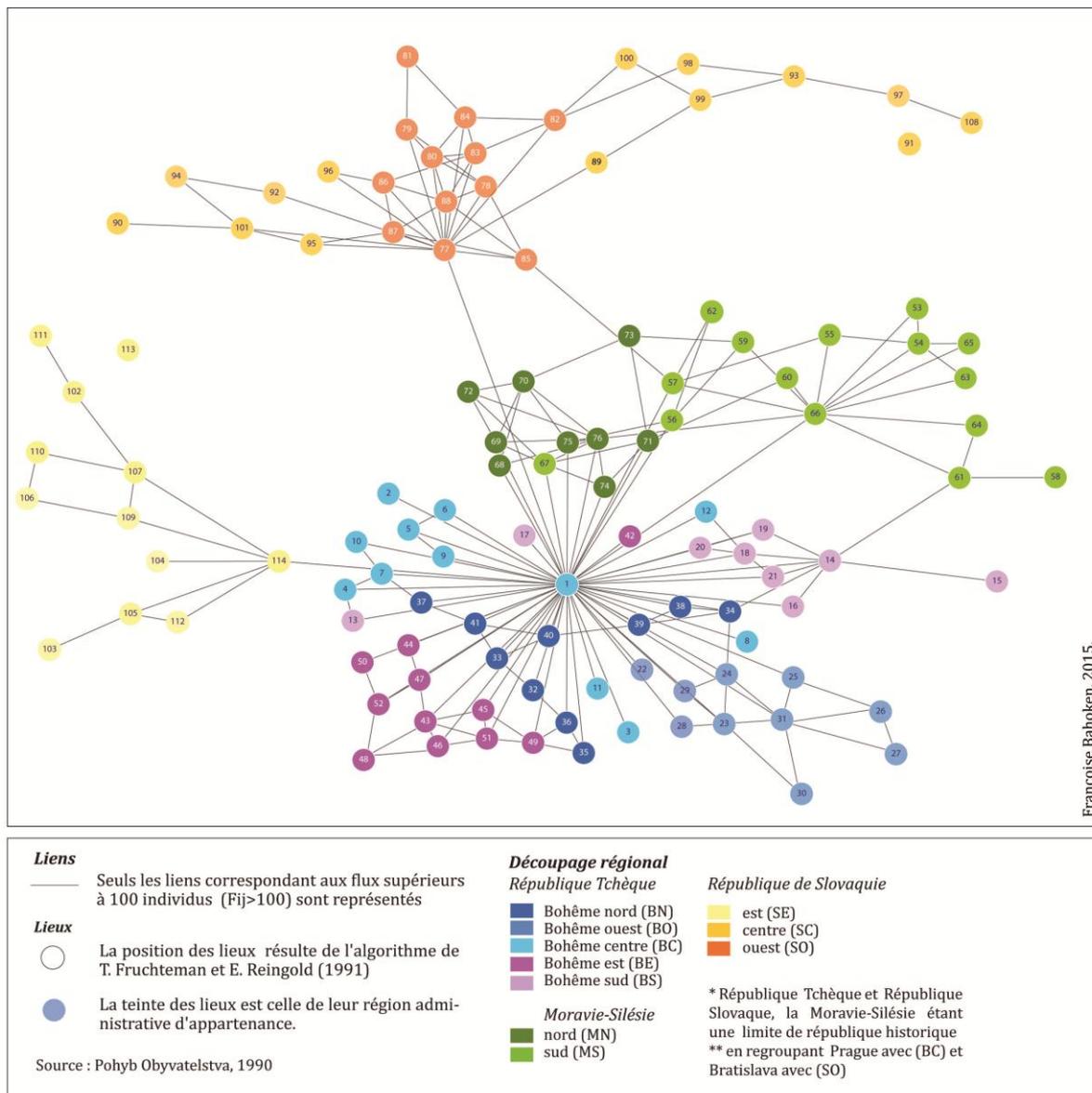
La seconde approche tient dans la sélection des liaisons à représenter, dans le filtrage des données (liens ou nœuds), souvent associé à un algorithme de repositionnement des lieux de manière à minimiser le chevauchement des liens. Plusieurs algorithmes de placement des nœuds sont disponibles dans la littérature, une présentation est proposée dans (Bahoken et al., 2011 : 5-8).

La Figure 2-9 présente un exemple de cette approche appliquée au graphe des flux de l'ex-Tchécoslovaquie (1990) : seule la sélection des flux dont la valeur est strictement supérieure à 100⁵⁹ individus est représentée, décrivant ainsi 373 liaisons qui correspondent à 0,03% de l'ensemble représenté sur la Figure 2-8. La position des nœuds dans ce graphe n'est pas aléatoire, elle résulte de l'application de l'algorithme de T. Fruchterman et E. Reingold (1991), supposé bien adapté à la visualisation des communautés spécifiques aux grands graphes (nous y reviendrons dans la section

⁵⁹ A ce stade, il n'est pas nécessaire de préciser les raisons du choix de ce critère, la procédure de sélection des valeurs de flux à cartographier fait l'objet des chapitres 7, 8 et 9.

2.3.3.2). La couleur des points ne l'est pas non plus, elle fait référence à la région d'appartenance du district dans le découpage territorial de l'ex-Tchécoslovaquie de 1990 présenté sur la Figure 2-3, page 59.

Figure 2-9. Le graphe des relations interdistricts de l'ex-Tchécoslovaquie (1990)

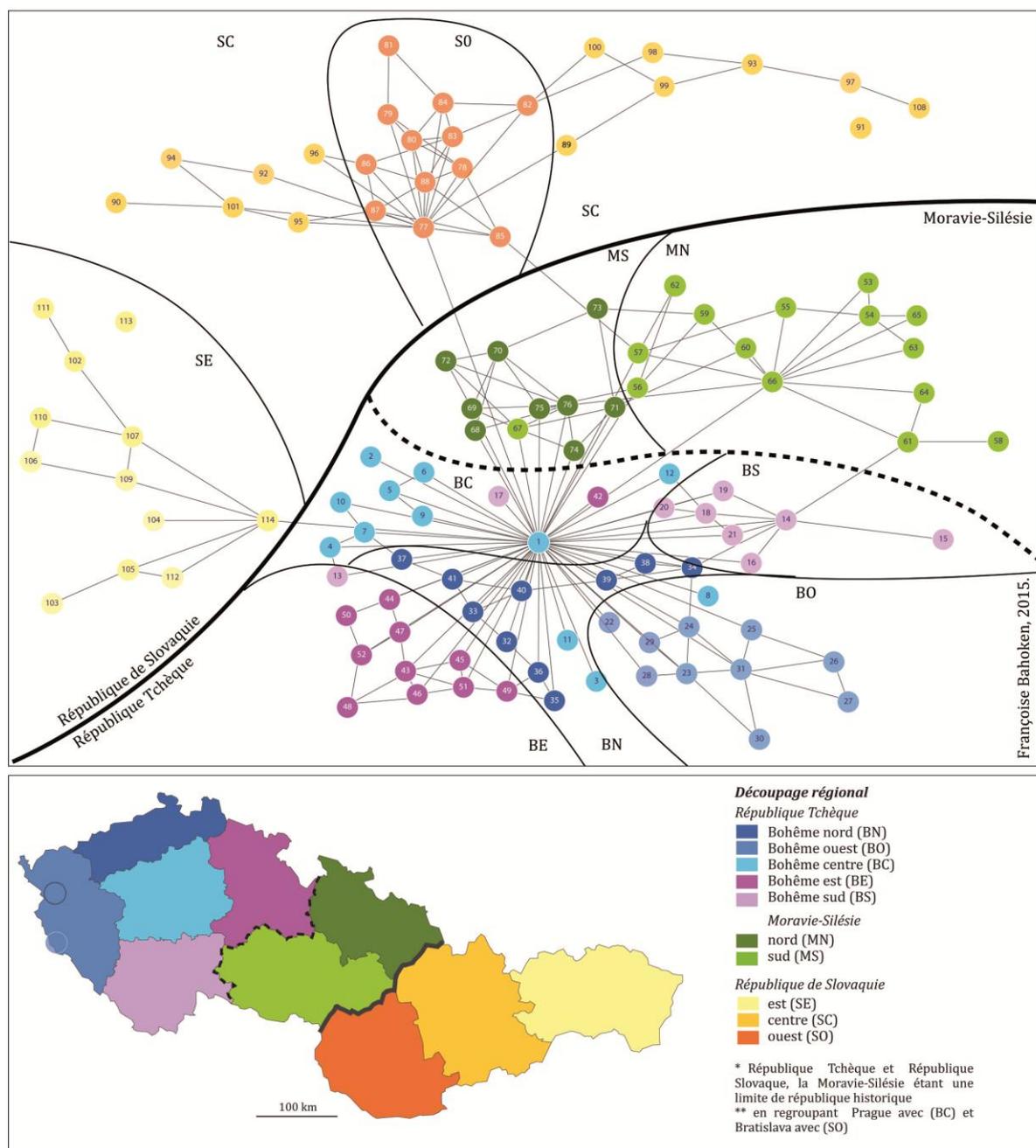


Ce graphe des relations est typique de la représentation d'un diagramme lien-nœud, combinant sur la même figure la représentation des liens et des lieux.

Son efficacité est avérée dans la mesure où la structure territoriale de ces flux est bien visible sur le graphe. En effet, si on la compare avec la cartographie des régions, on s'aperçoit que les communautés ainsi formées correspondent aux limites régionales, elles-mêmes faisant écho à celles des anciennes républiques Tchèques et Slovaque.

La Figure 2-10 compare le graphe de la Figure 2-9 avec le découpage régional.

Figure 2-10. Le graphe des relations interdistricts de l'ex-Tchécoslovaquie (1990) – variante



2.2.2 Le principe

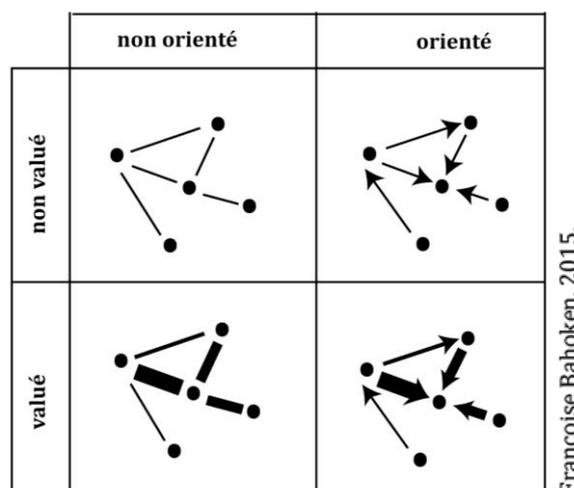
Le principe de réalisation d'un graphe consiste à projeter sur un plan la structure des relations qui existent entre des sommets, en positionnant ces derniers de manière à limiter le chevauchement des liens.

Plusieurs méthodes peuvent être mises en œuvre pour tracer un graphe. Le cas général étant réalisé sur un plan en deux dimensions. La figure est composée d'arcs qui relient des sommets représentés par des symboles, généralement des points.

Un graphe planaire est ainsi la figure réalisée sans qu'aucun arc ne se croise à l'inverse des graphes non planaires, caractérisés par le chevauchement des liens. Trois principaux types de graphes peuvent ainsi être réalisés, « [...] en fonction de la nature des liens existants : – les graphes binaires ou booléens (binary) où les liens traduisent l'existence d'une relation entre deux sommets ; – les graphes

valués (*valued*) où les liens traduisent la présence et l'intensité de la relation ; – les graphes signés (*signed*) où les liens expriment soit l'attraction, soit le rejet [...] (Beauguitte, 2010)⁶⁰. La Figure 2-11 illustre les quatre différents types de graphes envisageables, issus d'une même matrice : les graphes non valués (binaires) orientés et non orientés, ainsi que les graphes valués orientés et non orientés.

Figure 2-11. Les différents types de graphes



Françoise Bahoken, 2015

Dans un graphe, la position des nœuds et des liens dans le plan peut être arbitraire ou non, comme dans le cas de la Figure 2-8 où nous avons utilisé l'algorithme de Fructermann et Reingold (1991) pour placer les lieux, mais il existe une multitude de solutions de placement. Dans la pratique, le positionnement des nœuds suit généralement un ordonnancement particulier de manière à illustrer une forme géométrique donnée : un cercle ou un polyèdre, par exemple, ou satisfaire à des critères esthétiques propres à l'auteur, théoriques ou thématiques. Dans ce cas, le dessin du graphe est soumis à la prise en compte de contraintes de différentes natures qu'il n'est pas toujours possible d'articuler au sein d'une même figure.

2.2.3 Les contraintes

La particularité de la représentation des matrices sous la forme de graphe est d'intégrer au sein d'une même figure, un certain nombre de contraintes⁶¹ à la représentation des N nœuds et des L liens, dans un espace qui n'est pas forcément métrique. Ces contraintes sont conceptuelles, esthétiques ou encore sémiologiques. Cependant, seules les trois premières seront présentées dans le cadre de cette section : les contraintes d'ordre sémiologique, compte tenu de leur importance en cartographie des flux, font l'objet de la seconde section du chapitre 4.

2.2.3.1 Les contraintes conceptuelles

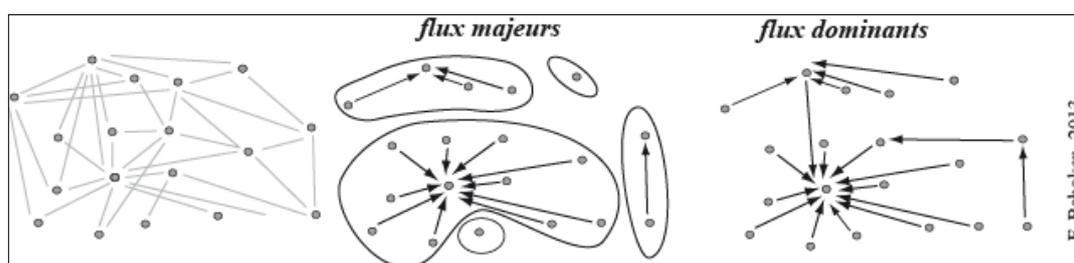
Les contraintes conceptuelles sont, en général, appliquées en amont de l'étape de représentation graphique (Bunge, 1962, 1966) et cela, quel que soit le type de graphique à réaliser. Elles portent sur le type de données ; sur les aspects scalaires qui correspondent au niveau d'observation et par conséquent, à l'échelle de la représentation ; enfin sur l'importance accordée aux nœuds.

⁶⁰ Cette définition des liens ne fait référence qu'aux seules données des matrices numériques d'adjacence où les modalités des valeurs des cellules (i,j) sont binaires (1,0) et la localisation des sommets dans le plan est inexistante. Nous y reviendrons dans la section 3 du chapitre.

⁶¹ Le site internet <http://www.visualcomplexity.com> [lien vérifié le 02/08/2015] est un bon exemple de la variété et de la complexité des procédés de représentation graphique sous la forme de graphe.

La contrainte du type de données est liée à leur format, selon qu'elles sont binaires, discrètes ou continues. En effet, le type de données influe sur les traitements réalisés en amont de la représentation, par conséquent sur la manière de représenter les résultats obtenus. Pour le démontrer, considérons le cas de la représentation classique de flux majeurs et dominants d'un graphe quelconque, tels qu'ils sont figurés par Puebla (1987), cité par C. Ducruet (2011). Les deux images présentées sur la Figure 2-12 décrivent, à partir du graphe (partie gauche), les flux majeurs (partie centrale) qui correspondent aux flux les plus grands émis par chaque lieu (les flux de rang 1 ou les plus gros), et les flux dominants (partie de droite). Nous ne rentrerons pas dans les détails de l'explicitation de la procédure de sélection des flux dits majeurs et dominants, celle-ci faisant l'objet du chapitre 9.

Figure 2-12. Exemple de représentation des flux majeurs et dominants

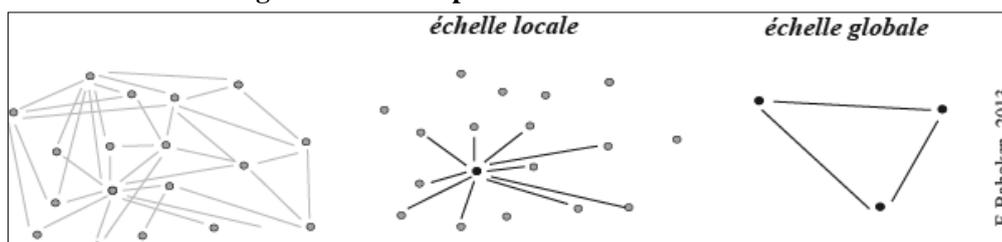


Source : D'après Puebla (1987) cité par C. Ducruet, 2011.

Le graphe représentant l'ensemble des relations observées, situé sur la partie gauche de la Figure 2-12, ne met en évidence aucun motif particulier, aucune structure spécifique à l'inverse de celui des flux majeurs, situé au centre, qui permet d'isoler différentes composantes connexes, correspondant à des communautés bien identifiées. Le graphe des flux dominants conduit à l'inverse à une représentation plus hiérarchique.

Le second type de contraintes concerne les aspects scalaires, au sens de niveau d'observation des relations entre les nœuds : à l'échelle globale de l'ensemble de la matrice ou à l'échelle locale d'un ou de plusieurs nœuds en particulier. La figure peut ainsi consister à ne représenter que les liens entrants ou sortants d'un lieu en particulier, comme illustré par la partie centrale de la Figure 2-13, qui illustre l'échelle locale. Cette contrainte scalaire porte également sur la résolution (ou l'échelle) de la matrice. En effet, si l'on compare la partie droite de la Figure 2-12 à celle de la Figure 2-13, on observe que le changement d'échelle consiste non plus à raisonner au niveau individuel (local) mais au niveau agrégé de groupes de nœuds formés autour de nœuds particuliers (ici, en l'occurrence dominants), visibles à un autre niveau hiérarchique. La contrainte scalaire est définie au regard du contexte de l'analyse, selon qu'elle est menée au niveau du lieu ou de l'ensemble.

Figure 2-13. Exemple de contraintes scalaires



Le troisième type de contrainte conceptuelle porte sur la prise en compte éventuelle d'un référentiel lié à la position des nœuds (en x et y) dans le plan. Elle interroge la question fondamentale du choix de la

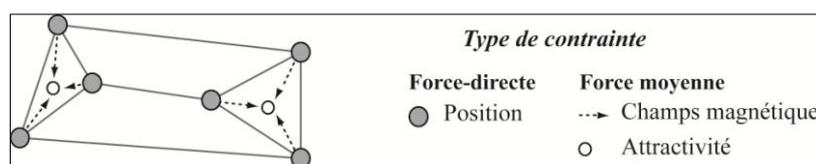
méthode de placement des liens et des nœuds d'un graphe non planaire dans un plan. Les contraintes de type « *force-directed* » (« *forced-based* ») renvoient à l'énergie (*forced*) et à la contrainte (*directed*) exercée sur un nœud. Utilisant une méthode issue de la physique, de minimisation de l'énergie mesurée par l'espacement entre sommets, elles peuvent être mises en œuvre à l'aide d'une variété d'algorithmes visant généralement à rapprocher les lieux reliés et à limiter les croisements des liens. Compte tenu de la difficulté de traduire ce terme, nous faisons le choix d'utiliser ce terme anglais.

2.2.3.2 Les contraintes de « *force-directed* »

L'application de contraintes, telles que celles de « *force-directed* », que nous prenons pour exemple consiste à optimiser le placement des nœuds et des liens pour améliorer la lisibilité du graphe. Leur objectif est de les repositionner pour minimiser le chevauchement des liens, puis de faire émerger des sous-groupes (des communautés ou sous-composantes du graphe) sous une forme visuelle particulière, en particulier lorsque la matrice est dense.

Apposer une contrainte de type *forced-directed* à un graphe non planaire consiste à forcer le *placement*, c'est-à-dire la position des nœuds ou des liens dans le plan, en appliquant d'une part, une *fonction d'énergie* et d'autre part, une pondération aux mesures des sommets, qui correspond à une forme d'attractivité des nœuds. Le choix d'un algorithme de positionnement n'est donc nullement anodin, car il pèse sur l'interprétation. Or il existe de très nombreuses variantes et alternatives à cet algorithme de force directe, l'une d'entre elles est illustrée par la Figure 2-14.

Figure 2-14. Exemple d'application de forces directes moyennes et d'attracteurs



Source : D'après, Cruz et Tamassia, n.d.

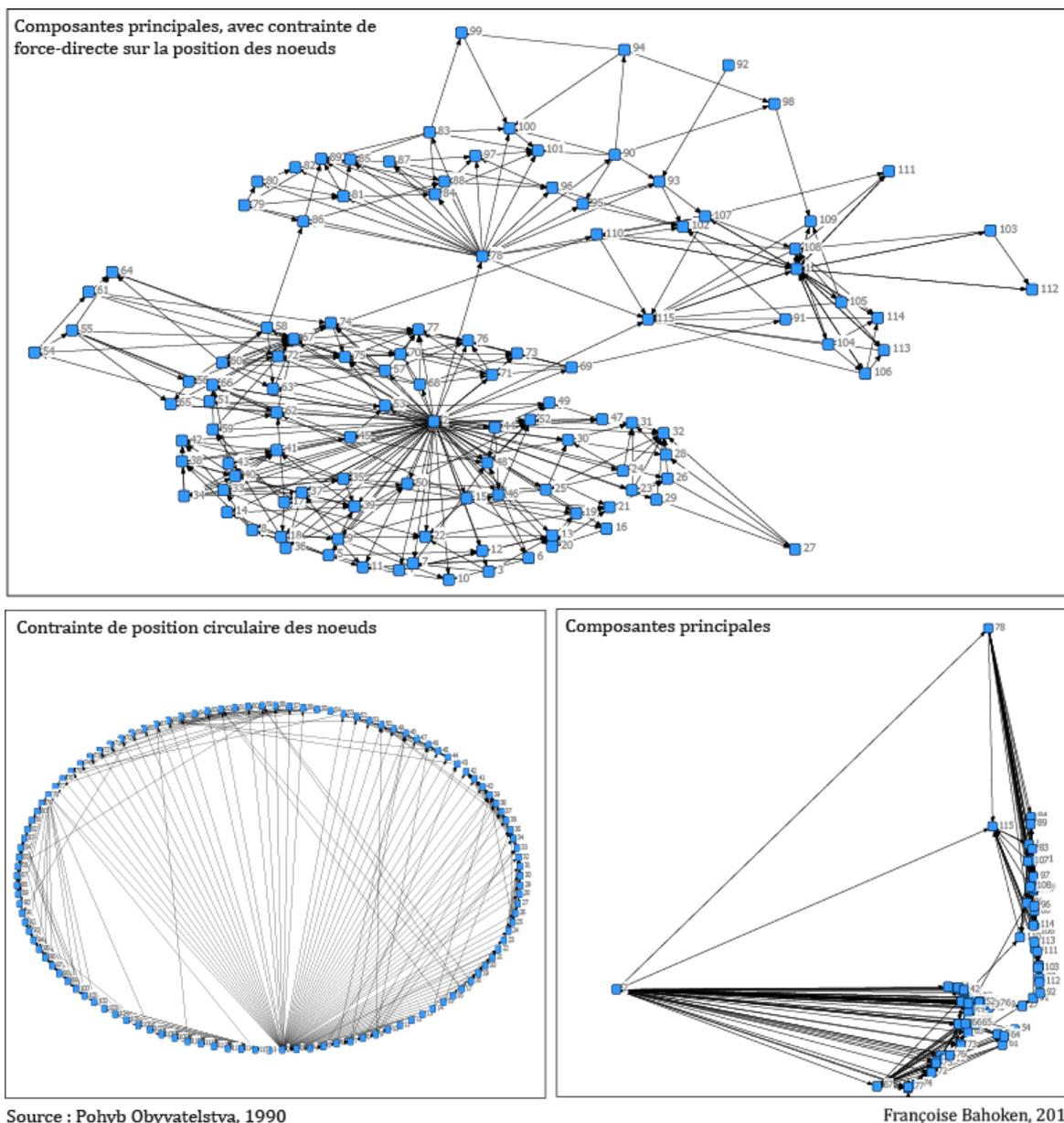
La Figure 2-15, située page suivante, présente différents dessins de graphe des relations interdistricts de l'ex-Tchécoslovaquie (1990) qui sont supérieures à 100. Il s'agit donc de variantes de la Figure 2-9 déjà présentées (page 69).

Sur la Figure 2-15, la différence entre les images présentées réside dans le type de contrainte appliquée à la position des lieux. Les représentations dites en « *composantes principales* » font référence à la notion d'équivalence structurelle⁶² appliquée à la position relative des nœuds les uns par rapport aux autres dans le plan. Elles décrivent ici deux voire trois composantes fortement connexes qui caractérisent l'appartenance des districts à l'une des deux républiques (Tchèque ou de Slovaquie). Ainsi, le fait que Prague (district n°1) s'oppose à tous les autres (voir Figure 2-9) s'explique alors logiquement par le fait que c'est le seul district à ne pas avoir de relation forte avec ... Prague ! Tous les autres districts présentent, en effet, un point commun : celui d'avoir une relation forte avec la capitale.

Quant à la représentation du même graphe sous une forme circulaire, elle n'a éventuellement un intérêt que sur le plan esthétique (la position des lieux ne répondant à aucun critère thématique).

⁶² Une équivalence structurelle existe entre deux sommets A et B si leurs voisinages respectifs (formé par l'ensemble des lieux avec lesquels ils sont liés directement) sont identiques, à l'exception des relations qu'ils ont entre eux. En d'autres termes, A et B sont structurellement équivalents s'ils sont en relation avec les mêmes sommets.

Figure 2-15. Exemples de graphes des relations interdistricts de l'ex-Tchécoslovaquie supérieures à 100, contraintes par la position des nœuds



Herman et al. (2000) ont proposé une classification de l'ensemble des contraintes de placement des sommets en trois catégories que sont : le « *ghosting* », qui fait disparaître – de l'affichage – les nœuds ainsi considérés comme des fantômes ; le « *hiding* » qui consiste à sélectionner des nœuds et le « *grouping* » (aussi nommé « *merging* ») qui consiste à agréger les nœuds. Ces solutions visuelles permettent, en réalité, de limiter le nombre de nœuds visibles, par conséquent le nombre de liens, ce qui a pour effet d'améliorer la lisibilité globale du graphique.

Les deux algorithmes « *forced-based* » les plus courants sont ceux de T. Kamada et S. Kawai (1989) et de T. Fruchterman et E. Reingold (1991). Le premier consiste, d'après I. Cruz et R. Tamassia à placer les nœuds les uns par rapport aux autres de manière à ce que leurs séparations respectent une mesure de distance calculée sur le graphe. Cette distance peut correspondre à une métrique réelle, la distance euclidienne, par exemple, comme dans le cas de l'application d'un facteur d'échelle ou non. Une option de cet algorithme permet de séparer les différentes composantes du graphe, afin de faire

émerger des communautés qui présentent des relations plus importantes entre elles qu'avec les autres. L'algorithme de T. Fruchterman et E. Reingold (1991) est, quant à lui, fondé sur des caractéristiques nodales d'attractivité ou de répulsion, ainsi que sur une fonction de type *masse-ressort* appliquée aux liens. Etabli en analogie avec le principe gravitaire, il permet également de distinguer des sous-groupes et semble, en pratique, plus adapté aux grands graphes étant par ailleurs orientés. C'est pour cela que nous mobiliserons ce dernier en cas de besoin.

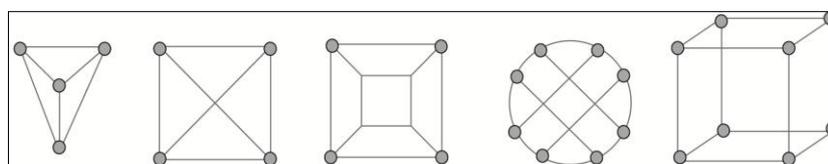
A la méthode de placement optimal des nœuds à l'aide d'une contrainte est généralement associée une contrainte de *minimisation des chevauchements* qui consiste à gérer leurs croisements. Celle-ci consiste à appliquer des angles au tracé des liens – qui deviennent ainsi des courbes de Bézières – afin d'optimiser l'affichage des liens. La nécessité d'appliquer une contrainte de position des sommets du graphe, pour minimiser les chevauchements des liens, traduit le besoin d'améliorer la lisibilité de la figure. Ces contraintes n'opèrent toutefois que sur les aspects liés à la visualisation des relations, tout comme celles qui ont pour objectif de représenter l'orientation de la relation, en illustrant *l'effet de convergence*. Nous reviendrons sur ces points particuliers dans le chapitre 4.

Des contraintes esthétiques sont généralement appliquées dans un second temps. Elles n'agissent que sur l'apparence de la figure, sur le *design* du graphe. Mais elles peuvent aussi modifier le sens de la figure résultante.

2.2.3.3 Les contraintes esthétiques

Les contraintes esthétiques portent sur le style et le *design* de la figure. Elles sont généralement associées aux contraintes précédentes de forces-directes qui agissent sur la position des nœuds. Elles consistent à donner l'apparence d'une structure particulière au graphe, circulaire, par exemple. Les plus importantes sont liées à la volonté de *placement optimal* des nœuds afin de garantir d'une part, la visibilité des liens et de l'autre, la structure d'ensemble du système selon un modèle particulier. D'après G. Di Battista et al. (2000), l'application d'une contrainte esthétique consiste, par exemple, à i) placer les nœuds les plus connectés au centre de la figure afin de minimiser les croisements entre les liens ; ii) à les forcer à se placer sur les bords de la figure, afin de donner l'impression d'une symétrie ; iii) ou encore à les disposer dans un ordre particulier afin d'obtenir des formes circulaires, orthogonales et volumiques. La Figure 2-16 illustre ces différentes possibilités.

Figure 2-16. Exemples de contraintes esthétiques de placement optimal des nœuds



Source : D'après, Cruz et Tamassia, n-d.

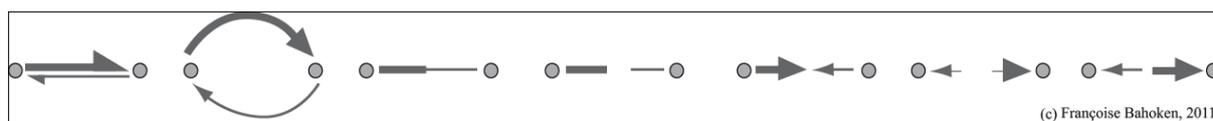
Il est très difficile, en pratique, de satisfaire à plus d'un de ces critères esthétiques en raison de verrous techniques et parfois technologique. Le style de la figure étant souvent le résultat de compromis réalisés dans le choix de contraintes à satisfaire compte tenu des moyens dont on dispose, sachant qu'il est nécessaire de considérer la sémiologie graphique, ce qui n'est généralement pas le cas.

2.2.3.4 Les contraintes sémiologiques

La sémiologie graphique permet de résoudre certaines contraintes posées par la représentation graphiques, notamment celles qui permettent d'aboutir à une figure ayant un sens. En effet, elle permet

de choisir judicieusement le procédé graphique à utiliser en fonction du type de liaisons (flèche, tête de flèche, lien, bande ...), autrement dit du type de matrice. Ainsi, lorsque les liens sont *asymétriques*, par exemple, le recours à la sémiologie graphique est un bon exemple d'amélioration des graphes produits avec des algorithmes qui n'en tiennent pas compte (nous y reviendrons dans la troisième section du chapitre 3). Dans ce cas particulier, il est également nécessaire de recourir à la flèche et de placer les liens de manière à illustrer *l'effet de convergence* ou de divergence vers ou depuis les nœuds, comme illustré par la Figure 2-17.

Figure 2-17. Exemples de contraintes sémiologiques sur le type de liaisons



Les contraintes liées à la sémiologie graphique (Bertin, 1967) portent sur le choix des variables visuelles à appliquer à la représentation des nœuds et des liens dans le plan : leur taille, leur forme, leur couleur, leur valeur et leur orientation. Elles s'appliquent à tous les types de graphiques aussi nous n'allons pas les présenter de manière extensive ici. Une synthèse récente de la manière dont les aspects sémiologiques sont pris en compte dans le dessin de graphe est présentée dans une synthèse du groupe fmr (Bahoken et al., 2013).

L'application de ces contraintes sémiologiques, dans le contexte spécifique de la cartographie de flux, faisant de la seconde section du chapitre 4, nous présentons maintenant les variantes au modèle du graphe.

2.2.4 Les variantes

Les variantes au modèle du *graphe* peuvent être appréhendées comme le résultat de l'apposition de contraintes qui sont, comme nous l'avons vu, de différents ordres. Celles qui conduisent à une variation de la figure sont essentiellement les contraintes des algorithmes de placement des sommets qui agissent sur le *design* de la figure. La prise en compte simultanée de plusieurs contraintes liées à la position des nœuds, à l'esthétique et la sémiologie graphique, est actuellement un axe majeur des développements réalisés dans le champ de la visualisation graphique des données relationnelles. Ce champ de la recherche est si fécond qu'il n'est pas envisageable de présenter ici toutes les propositions, lesquelles nous éloigneraient d'ailleurs de notre objet d'étude.

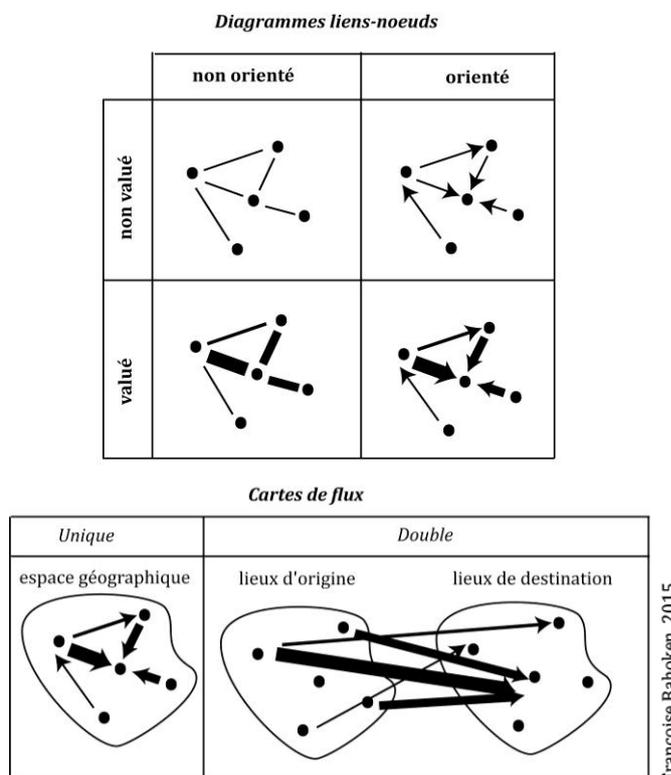
Le lecteur intéressé par les développements récents de nouvelles formes de visualisation pourra consulter, par exemple, les productions du groupe de recherche Aviz⁶³ de l'Institut national de recherche en informatique et en automatique (Inria) qui sont, en ce sens, exemplaires. Deux directions sont notables dans leurs travaux : la première consiste à créer des modèles hybrides qui combinent les modèles matriciels et du graphe au sein d'une même figure afin de démultiplier les possibilités d'exploration des données, comme dans le cas de NodeTrix © (Henry et al., 2007 ; Henry, Fekete, 2008 ; Fekete, 2011) également visible chez A. Rae (2009), la seconde à proposer de nouveaux procédés, tels que GeneaQuilts© (Fekete, 2010) présenté sur la Figure 2-18.

⁶³ Ces travaux ont été récemment récompensés, une vue synoptique est présentée sous la forme d'un poster accessible à l'adresse : <http://www.aviz.fr/wiki/uploads/Main/AVIZ-Poster-2012.pdf> [lien vérifié le 02/08/2015].

proposée par L. Beauguitte (2010), en y ajoutant un niveau correspondant à la spatialisation éventuelle de ses entités comme composante intrinsèque au nœud.

La Figure 2-19 présente les deux familles de graphes non planaires, selon qu'elles sont spatialisées (comme dans le cas de la carte) ou représentées dans un espace quelconque.

Figure 2-19. Les deux familles de graphes non planaires



Ainsi, aux quatre types de graphes proposés par l'auteur (binaires orientés ou non ; valués, orientés ou non) dont la particularité est de ne pas nécessairement tenir compte de l'espace, doivent être ajoutés les cartes de flux pour lesquelles l'espace est essentiel. On distingue alors la carte de flux ordinaire, où l'espace est représenté sur un fond de carte « *unique* » (Tobler, 1987) et la double carte de flux qui conduit au dédoublement de l'espace, où les lieux d'origine sont représentés séparément des lieux de destination.

Sur la Figure 2-19, la carte unique correspond à la version classique où les lieux d'origine et de destination sont placés sur le même fond de carte. Celle-ci présente différentes variantes selon que les liens et les lieux sont représentés conjointement ou non (dans ce dernier cas, on compte, par exemple, la carte de flux en points proportionnels, qui conduit à un changement de paradigme de la représentation puisque les flux y sont représentés du seuls point de vue des lieux) ; ou encore selon que la position des lieux correspond à une métrique géographique ou non (à l'inverse des cartes en anamorphoses scalaires ou vectorielles). La « double » carte de flux est dérivée du modèle des coordonnées parallèles où les lieux d'origine et de destination sont placés sur des axes (ou des fonds de carte) distincts.

Avant de présenter ces différentes variantes de la carte de flux, il est nécessaire de définir l'objet « Carte de flux ». A la différence des précédents modèles de représentations graphiques des structures relationnelles, sous la forme de matrice ou de graphe, la carte est une figure nécessairement spatialisée. « *Outil d'exploration par excellence, la carte possède cette qualité unique d'offrir à son*

lecteur un référentiel spatial, sinon maîtrisé, tout au moins plus familier que l'espace abstrait des autres types de graphiques statistiques. Cette qualité inhérente à son objet même (l'espace), peut être grandement amplifiée, jusqu'à permettre une véritable "déambulation" spatiale. » (Banos, 2001). Ce n'est donc pas seulement l'exploration et la visualisation graphique de données disponibles dans une matrice de flux qui intéresse la cartographie, mais aussi la visualisation simultanée de la structure spatialisée du phénomène (en l'occurrence des flux) et de ses caractéristiques. La carte autorise l'exploration « des déambulations » dans l'espace parce que, à l'inverse du graphe, elle projette la structure des relations entre les différentes entités, dans un espace géographique, c'est-à-dire dans un plan assorti d'une métrique (en l'occurrence géographique).

La différence entre le graphe et la carte de flux réside, par conséquent, dans la manière dont l'espace est pris en compte dans la représentation pour symboliser des relations entre des objets localisés dans un espace : **le graphe permet de visualiser la structure des flux là où la carte de flux permet de projeter cette structure dans un espace métrique, généralement géographique. Mobiliser une carte de flux, plutôt qu'un graphe, traduit alors l'importance accordée à la situation des lieux dans cet espace, celle de leurs espacements.**

2.3.1 L'objectif

Au sens général, « *La carte est une représentation géométrique plane, simplifiée et conventionnelle de toute ou partie de la surface terrestre, et cela dans un rapport de similitude convenable que l'on appelle l'échelle* »⁶⁴ (Joly, 1976 : 1). Cette définition fait référence à la procédure d'acquisition et de restitution des formes terrestres d'occupation des sols, à l'origine des bases de données d'information géographique disponibles sous la forme de référentiels nationaux tels que la base de données topographique (BDTOPO©) de l'Institut national de l'information géographique et forestière (IGN) par exemple. Notre recherche ne se situe pas à ce moment là de la production cartographique, mais plutôt dans une phase aval évoquée dans la définition plus ancienne du géographe russe K. A. Salitchev (1967), cité par M. Beguin et D. Pumain (1994, 2003). Du point de vue de l'auteur, « *La carte est une représentation réduite, généralisée, mathématiquement précise de la surface terrestre sur un plan montrant la situation, la distribution et les rapports de divers phénomènes naturels et sociaux, choisis et définis en fonction de chaque carte. La carte permet également de montrer les variations et les développements des phénomènes dans le temps ainsi que leurs facteurs de déplacement dans l'espace* ». (Salitchev, 1967, cité par Beguin, Pumain, 1994 : 5).

Cette seconde définition présente plusieurs intérêts dans le cadre de notre recherche. Le premier est celui d'avoir, comme le soulignent M. Beguin et D. Pumain (*ibid.*), une finalité géographique. La carte est nécessairement réalisée sur un plan assorti d'une métrique qui correspond à un système de projection cartographique, que la figure fasse apparaître – ou non – le maillage spatial sous-jacent : la carte de flux ayant pour objectif de représenter le transfert d'une quantité d'un point à un autre dans un espace.

2.3.1.1 Représenter le transfert dans l'espace

La définition précitée de K. A. Salitchev présente l'avantage d'intégrer les concepts de déplacement dans l'espace, c'est-à-dire le changement de localisation ou de position des objets ainsi que la distance parcourue. Le premier des objectifs d'une carte de flux est effectivement de représenter le transfert

⁶⁴ Le Comité français de cartographie (CFC) propose une définition similaire de la carte, en incluant la manière dont les motifs sont représentés et donnés à voir au lecteur : elle correspond à une « [...] *représentation géométrique conventionnelle, généralement plane, en positions relatives, de phénomènes concrets ou abstraits, localisables dans l'espace ; c'est aussi un document portant cette représentation ou une partie de cette représentation sous forme d'une figure manuscrite, imprimée ou réalisée par tout autre moyen.* » (CFC, 1990).

dans l'espace d'un processus de mobilité spatiale, s'inscrivant *a fortiori* dans un espace géographique donné.

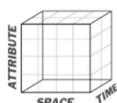
La mobilité spatiale caractérise le déplacement d'une entité dans l'espace et dans le temps qui constituent deux cadres d'analyse distincts mais imbriqués, au même titre que les catégories thématiques (personnes, biens) et leurs caractéristiques (respectivement le genre et le type de produit). Elle met en relation au moins deux lieux, dans le cadre d'un déplacement qui est observé soit au niveau individuel, soit à un niveau agrégé comme nous l'avons présentée dans le chapitre 1. Dans le premier cas, le déplacement prend la forme d'un cheminement, d'un chapelet de lieux qui correspond, par exemple, à l'ensemble des lieux pratiqués par l'individu au cours de la journée, de la semaine ou de sa vie. Dans le second cas, le déplacement est celui d'un agrégat, il correspond au flux.

La carte de flux vise par conséquent à représenter la quantité qui s'est déplacée entre des couples de lieux. Elle peut donner lieu, *a minima* et à l'instar du graphe, à la représentation de l'existence d'un lien entre les mêmes couples de lieux. Le déplacement intervenant nécessairement dans un intervalle de temps, la carte de flux intègre souvent implicitement la composante temporelle dans sa représentation. La temporalité du déplacement est, d'ailleurs, souvent comprise dans la valeur du flux lui-même. Etant donné que nous reviendrons longuement sur les aspects liés à la composante spatiale (chapitre 6 et 7), nous consacrons les sections suivantes à la prise en compte du temps dans la cartographie de flux.

2.3.1.2 Représenter la structure du changement ...

La représentation du changement dans l'espace nécessite généralement la prise en compte du moment ou de l'intervalle de temps du phénomène (voir Encart 2-2). La carte de flux intègre, dans la majorité des cas, le temps de manière implicite.

Encart 2-2. Sur la prise en compte du temps en cartographie



En cartographie, la prise en compte de la dimension temporelle peut être envisagée de plusieurs manières : – du point de vue du processus d'édition : « *les cartes sont comme le lait : l'information est une denrée périssable et il est prudent de vérifier sa date de validité* »

(Monmonnier, 1993 : 91) ; – de la date de collecte des données ; – de la date de conception, de création, d'impression et d'actualisation de la carte, mais également – du point de vue des données et de leur représentation (lorsqu'elle est animée). Tenir compte du temps des données est essentiel dans la compréhension des dynamiques spatiales. En géographie, le qualificatif de *spatio-temporel* associe inévitablement l'espace (x,y) au temps (t), notions qui sont toutes deux fondamentales. B. Berry (1964) a formalisé cette relation dans un bloc diagramme d'espace-temps où les objets spatio-temporels sont définis sur trois références (x, y, t) auxquelles il faut associer une composante thématique (z). Plusieurs auteurs ont proposé une analyse de la prise en compte du temps dans le processus de cartographie. C'est M. Harrower qui présente différentes manières d'intégrer le temps, dans un article consacré à la prise en compte du temps dans les cartes (Harrower, 2007 : 4). P. Haggett (1990) avait auparavant suggéré que le changement temporel pouvait être représenté de quatre manières : comme une « *tendance* », sous une forme « *cyclique* », « *récurrente* » ou « *transitive* ». La typologie en cinq classes proposée par Vasiliev (1997) quelques années plus tard fait état de « *la date du phénomène* » aussi qualifié de « *moment* », de « *l'intervalle* » ou « *durée* », de « *temps structuré* », de temps comme « *une mesure de la distance* » et de « *temps assimilé à une horloge* ».

En effet, tous les changements se manifestent dans un temps donné et sur une période. D'après G. Palsky (2004), « [...] *la carte de flux en est l'expression la plus simple* [du temps], *car la variable*

temps n'intervient que comme durée de totalisation (flux migratoire annuel par exemple) » (Palsky, 2004a : 348-349). L'auteur s'interroge, par ailleurs, sur les raisons qui ont poussé J. Bertin à classer le temps comme variable ordonnée alors qu'il s'agit bien d'une variable quantitative, mesurable ; il recense trois manières de représenter le changement là où R. Brunet (1987) en propose quatre, parmi lesquelles celle qui consiste à utiliser le procédé de la flèche.

La carte de flux ne saurait pour autant être réduite à la seule représentation du changement spatial ou du déplacement d'objets : du transfert dans l'espace, dans un sens ou dans un autre, de quantités. D'après (Dorling, 1991 : 104-105), la carte de flux est supposée représenter de manière statique⁶⁵ la structure du changement, son origine tout en mettant en évidence ce qui a changé. L'auteur précise que pour décrire le changement d'une structure sociale, à n'importe quel moment, un seul état suffit pour chacune des unités spatiales. Pour montrer où la structure a changé, il est nécessaire de disposer d'au moins deux états, pour chacune des unités spatiales. Pour présenter comment la structure a changé et ce qui s'est mû d'un endroit vers un autre, des informations complémentaires sur les relations entre les unités spatiales sont requises. Il s'agit généralement des flux mesurés, c'est-à-dire des valeurs observées mais cette information sur la quantité de flux transférée n'est pas si évidente à représenter, en particulier lorsqu'ils sont nombreux puisque la figure à l'apparence d'un « chaos » (Tobler, 1987). Il s'agit là d'une contrainte majeure de la cartographie de flux qui consiste à réduire la *complexité graphique* apparente de la figure : nous l'examinerons dans la section 2.3.3.1.

Comme l'indique G. Palsky, « [...] *le problème se complique encore lorsque le cartographe souhaite rendre compte d'un ensemble de mouvements* » (Palsky, 2004a : 350), c'est-à-dire pour les matrices denses dont on connaît le cheminement spatiotemporel (nous y reviendrons dans le chapitre 6), en raison de la manière dont le temps du phénomène est pris en compte dans la cartographie ce qui, d'après l'auteur, contribue à le déstructurer⁶⁶. N'oublions pas que la carte de flux, en tant que variante du graphe non planaire spatialisé, est de fait concernée par le second objectif qui préside à la réalisation du graphe, que nous avons déjà mentionné, à savoir : la réalisation d'une figure lisible.

Représenter le changement sur une carte de flux passe aussi par la représentation d'une différence. Autrement dit, montrer le changement consiste à donner à voir ce qui a changé, ce qui suppose de représenter une différence par rapport à un état initial. Cette représentation du changement, de la dynamique du phénomène passe donc invariablement soit par la mise en place d'un indicateur fondé sur le principe de la différence de valeurs, soit par la traduction graphique de ce changement. Nous reviendrons plus précisément sur ces aspects dans les chapitres 3 et 4 consacrés aux méthodes usuelles de cartographie des matrices de flux.

La cartographie des matrices de flux denses, qu'elles décrivent des relations – valuées ou non – entre des lieux, revêt par conséquent un premier niveau de complexité qui la distingue de la cartographie statistique de données scalaires : celui de réaliser une figure lisible. L'enjeu n'est pas le moindre puisque, comme l'indique G. Palsky (2004a) : « *Dans nos structures sociales, c'est [davantage] le mouvement entre les lieux, plutôt que les modifications qui se produisent en leur sein, qui sont responsables de leur structuration* » (*ibid.*). Pour cela, les solutions apportées passent l'application d'un critère de sélection, d'une réduction de l'univers ou de la résolution des données, pour produire

⁶⁵ A l'heure où D. Dorling (1991) rédige sa thèse, les possibilités de la représentation dynamique et interactive ne sont pas encore développées.

⁶⁶ L'auteur ajoute ceci : « *Ainsi, plutôt qu'une conception newtonienne du temps (un flux mesurable et quantifiable), la carte paraît transmettre une temporalité 'leibnizienne', conçue à travers l'ordre des événements, mais un ordre en tout état de cause partiel (valant pour chaque mobile en mouvement, mais non pour l'ensemble des mobiles). La carte déstructure le temps, en oblitérant sa linéarité et son irréversibilité. Où commence la lecture d'un système de mouvements ? Dans quel sens et dans quel ordre ? Seule la variation visuelle de troisième dimension peut permettre d'exprimer synchronisme et succession, à condition que la même variation affectant des mobiles différents corresponde à un moment précis* » (Palsky, 2004a : 350).

une figure lisible. Les valeurs de flux observées peuvent aussi être transformées en indices, permettant ainsi la représentation de gradients d'intensité des flux plus que de leur valeur brute, ce qui permet d'éviter le problème de la complexité graphique. Nous montrerons, dans le chapitre 5, qu'il n'est désormais plus obligatoire de sélectionner les données, voire de les réduire, pour pouvoir représenter de manière lisible, l'information contenue dans une matrice de flux observés.

2.3.1.3 ... et sa dynamique...

Cette capacité des cartes de flux à représenter les changements dans l'espace a contribué à sa dénomination de cartes dynamiques (Dorling, 1991 ; Dent 1999), en référence avec la représentation du mouvement qui est un phénomène non statique. Le caractère dynamique de la carte de flux peut aujourd'hui être appréhendé de diverses manières, selon que l'on s'intéresse aux aspects théoriques et conceptuels, aux questions méthodologiques ou bien aux aspects techniques.

Sur le plan théorique et conceptuel, la représentation de la dynamique des flux consiste, d'après W. Tobler (1979, 1982), à représenter une composante particulière des valeurs de flux obtenue dans le cadre d'une procédure de décomposition⁶⁷ de la matrice de données initiale. La représentation du caractère dynamique des flux est plus généralement appréhendée, en tenant compte du temps, dans l'objectif de comparer l'évolution des valeurs de flux ou d'indicateurs spécifiques des flux. Sur le plan thématique, le caractère dynamique des flux peut être mobilisé pour comparer le comportement de différents types de flux, c'est-à-dire de flux qui concernent des groupes d'individus différents ou des produits différents. La dynamique des cartes de flux peut, enfin, être appréhendée de manière technique, par le biais des outils permettant leur visualisation interactive et/ou animée : dans ce cas, la dynamique du support – de la représentation – se double alors à celle du phénomène. Elle nécessite alors d'articuler le temps intrinsèque des données, la temporalité du phénomène et le temps de la représentation, c'est-à-dire de l'animation – qui illustre la dynamique – lorsqu'elle est retranscrite sur un support écran.

Sur le plan technique, la prise en compte de la dynamique des flux ne se pose pas dans les mêmes termes, selon que la carte est disponible sous une forme dynamique⁶⁸ et interactive⁶⁹ ou non. Dans le premier cas, la dynamique est appréhendée sur le plan technique, en lien avec le support de la représentation tandis que dans le second, la dynamique est intégrée dans la représentation du flux, pour représenter le changement et donner l'illusion de la perception du mouvement associé à ce changement.

La question du type de support de la carte de flux est actuellement survalorisée. Elle apparaît aujourd'hui importante, en raison du développement des technologies de visualisation graphique liées à l'Internet, mais elle ne saurait être considérée comme fondamentale. En effet, les travaux du groupe de recherche *Cartomouv'* (Cartographie du changement, cartographie en mouvement), notamment les expérimentations développées dans le laboratoire de sémiologie ont permis de montrer la différence entre la représentation du changement dans l'espace et la perception du mouvement associé, liées au support écran. Ces deux composantes intrinsèques de la carte de flux (le changement de lieux et le mouvement lié au transfert dans l'espace) peuvent effectivement être révélées, lorsque la carte est proposée sous une forme dynamique. Dans ce cas, c'est davantage la dynamique de l'image qui est perçue, dans le cadre d'une simulation, que celle du phénomène (ce n'est pas tant le mouvement lié au

⁶⁷ Nous considérons cette procédure de décomposition au fondement de l'approche classique de cartographie des flux. Aussi une large place lui est consacrée dans le chapitre 3.

⁶⁸ La majorité des cartes de flux proposées sur un écran ne sont pas dynamiques, mais statiques : l'observateur n'ayant aucune possibilité d'interaction avec les données.

⁶⁹ L'interactivité est entendue au sens où l'utilisateur peut interroger une base de données à l'aide de requêtes et afficher le résultat de la sélection ainsi réalisée. Le contenu et/ou le fond, parfois la forme de l'affichage, sont modifiables par l'utilisateur.

flux qu'une représentation de celui-ci). Les développements informatiques ont largement contribué à augmenter les possibilités techniques d'interface entre l'observateur et la carte, en offrant de nouvelles perspectives de visualisation (Cheylan, 2007 ; Kaddouri, 2008), d'interaction et de simulation des processus représentés. Pour la carte de flux, cette extension des possibilités conduit à la possibilité de visualiser le mouvement géographique tel qu'il est intrinsèquement, c'est-à-dire en mouvement mais dans un rythme qui diffère nécessairement. Elle est bien illustrée par M.-J. Kraak (1998, 2003) qui a montré les avantages d'une géovisualisation dynamique des déplacements, par rapport à leur version statique, en mettant en scène dès 2003, la célèbre *Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813* de C.-J. Minard. Les pages internet⁷⁰ qui sont dédiées à cette expérimentation proposent plusieurs figures plus ou moins interactives qui conduisent, d'après l'auteur, à améliorer la perception des motifs représentés et à changer – effectivement – le rapport de l'observateur aux données.

La figure proposée sur un support dynamique est souvent accompagnée d'une série de variables spécifiques (un bouton radio ou un curseur, par exemple) permettant à l'utilisateur d'interagir avec la figure (Kraak, 1998 ; Zanin, Ben Rebah, 2011), de modifier son apparence et la quantité de figurés représentés. Cependant, ces aspects technologiques ne doivent intervenir, de notre point de vue, qu'à la marge du processus de création cartographique, les données devant être (correctement) analysées en amont de leur mise à disposition sur un écran. M. Baron et al. (2012) évoquent, à ce sujet, une « révolution » infographique qui conduit *in fine* à un « malaise » face à l'inadéquation entre l'effort de développement informatique réalisé et l'inefficacité cartographique d'un résultat qui ne demeure qu'au stade de la « curiosité cartographique ».

La réalisation d'une carte de flux étant, en réalité, d'abord liée à des considérations d'ordres théoriques et méthodologiques indépendantes du support, nous nous inscrivons dans une approche statique de la cartographie des matrices, que celle-ci conduise à une carte disponible sur un support papier ou numérique. Cela ne signifie pas notre désintérêt des aspects dynamiques ou interactifs du processus cartographique puisque nous avons participé au développement d'une application dédiée aux flux (Bonin et al, 2013), et fournissons en annexe une série de programmes informatiques devant conduire au développement d'une application dédiée à la sélection de flux à représenter. Nous signalons aussi la série d'applications⁷¹ développée par I. Boyandin et al. (2010, 2011) à savoir *Flowstrates* et *Jflowmap*, toutes deux développées dans un objectif d'exploration et de visualisation du changement temporel de la structure de flux entre deux voire plusieurs lieux. Ces applications présentent aussi l'avantage de combiner différents modes de représentation parmi lesquels celui de la matrice ordonnée, combiné ou non avec la « double carte » des flux. Elles ne sont toutefois pas uniques, c'est pourquoi nous en présentons une liste⁷² non exhaustive dans une section des références bibliographiques.

2.3.1.4 ... de manière agrégée

Lorsqu'elle est bien faite, la carte de flux donne à voir la structure d'ensemble d'un changement illustré par un système de relations agrégées entre des entités localisées dans un espace. Ce système de relations témoigne d'une interaction spatiale qui peut être observée à deux niveaux : au niveau individuel des acteurs ou des individus, dans une acception statistique et au niveau agrégé des lieux.

Le dictionnaire en ligne Hypergéo définit cette interaction spatiale comme « *une action réciproque (rétroaction) entre deux ou plusieurs acteurs ou lieux dans un système. Tous les échanges (de matière,*

⁷⁰ <http://www.itc.nl/personal/kraak/1812/> [Lien vérifié le 02/05/2015].

⁷¹ Une présentation vidéo est disponible à l'adresse suivante : <http://www.youtube.com/watch?v=UQPN7o6A3Cg&feature=youtu.be&hd=1> [Lien vérifié le 02/05/2015].

⁷² Une autre liste est également disponible sur une page du portail de curation thématique *Scoop.it* : <http://www.scoop.it/t/analyse-des-matrices-de-flux> [Lien vérifié le 02/05/2015].

de personnes, d'informations...), par ex. au niveau individuel entre producteur et client, entre partenaires, ou au niveau agrégé entre des villes ou des régions (ce sont les interactions), sont des interactions dans la mesure où ils provoquent des changements interdépendants dans les comportements ou dans les structures. Le terme a été introduit dans une théorie du mouvement par E. L. Ullman qui a posé comme condition nécessaire à l'interaction spatiale la complémentarité (compatibilité entre offre et demande) entre les lieux qui échangent et la transportabilité du produit (possibilité technique, existence d'infrastructures, coût supportable) [...]. Les interactions spatiales peuvent être utilement analysées à l'aide du modèle (qui fait partie plus généralement de la catégorie des modèles d'interaction spatiale) si elles sont fortement contraintes par la distance »⁷³.

Sur le plan théorique, le recours au concept d'interaction spatiale se justifie dans le cadre de lois générales d'organisation de l'espace par et pour les sociétés. Ces lois sont à la base d'une analyse spatialisée des phénomènes qui fournit un certain nombre d'hypothèses fondamentales de la géographie selon lesquelles ce qui se passe en un lieu n'est pas indifférent de ce qui se passe en d'autres lieux, à l'origine de l'autocorrélation spatiale. Certaines de ces lois découlent d'ailleurs directement de l'observation de la structure des flux. Le caractère systémique de l'organisation de l'espace par les sociétés qui les compose est ainsi à l'origine de la première loi de la géographie de W. Tobler (TFL), énoncée en 1970. D'après l'auteur, la TFL énoncée en référence au modèle gravitaire illustre le fait que les choses qui sont proches sont plus reliées entre elles que celles qui sont éloignées.

De la même façon que les interactions spatiales peuvent être observées au niveau individuel ou à un niveau agrégé, parfois aux deux niveaux simultanément, leur représentation cartographique est également envisageable au niveau des pérégrinations individuelles, sur une carte des déplacements, ou à un niveau agrégé dans une carte de flux. Il convient de noter que c'est la nécessité de retranscrire sur une carte les déplacements individuels qui est à l'origine de la cartographie de flux. La mesure des flux résulte d'ailleurs directement de celle des déplacements individuels, comme nous l'avons vu dans le chapitre 1.

2.3.1.5 ... mais aussi l'ampleur et l'inégalité de flux sélectionnés

La carte de flux a pour objectifs de représenter non seulement la structure du changement de manière agrégée mais aussi l'ampleur et l'inégalité des flux sélectionnés pour la représentation. Deux niveaux de représentation de ces flux sont également envisageables.

Le premier niveau de la cartographie de flux est celui de la « *représentation cartographique directe* » (Beguïn, Pumain, 1994 : 118) des interactions observées au sens où les données n'ont pas fait l'objet d'un traitement. Dans ce cadre, comme l'indique D. Poitevin (1999), la carte de flux permet de visualiser, « [...] *comme la carte de réseaux, les liaisons entre un certain nombre de points mais aussi l'ampleur de ces liaisons* » (Poitevin, 1999 : 147). Elle figure donc un ensemble de points, reliés par des liens qui correspondent à l'existence d'un flux. R. Brunet précise cette définition en introduisant deux éléments fondamentaux. Il souligne, en premier lieu, la différence entre la carte de flux (immatériels) et celle des réseaux (ou des flux matériels) de la manière suivante : « *Les cartes de flux ressemblent à celle des réseaux, mais s'en écartent souvent en raison de la complexité des relations à cartographier. [...] elles ont pour objectif de] montrer l'ampleur et l'inégalité des mouvements de biens, de personnes, ou d'informations entre un certain nombre de points* » (Brunet, 1987 : 90). L'auteur pose ensuite d'emblée le problème général de la cartographie des flux, dont on a vu qu'il concernait aussi le modèle du graphe : « *Au-delà d'un certain nombre de couples, la carte serait illisible, et l'on doit avoir recours à des artifices* » (*ibid.*)

⁷³ <http://www.hypergeo.eu/spip.php?article73> (texte rédigé par D. Pumain) [lien vérifié le 02/08/2015].

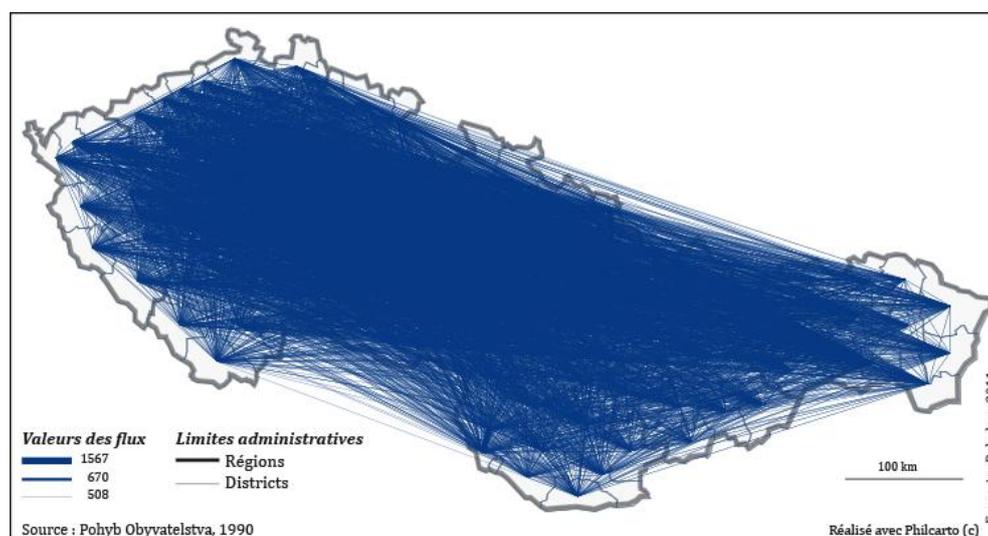
Le second niveau de la représentation cartographique des flux est celui des écarts des valeurs des flux à une loi générale. Le dictionnaire Hypergéométrie (*op.cit.*) précise pour ce faire qu'«*Un modèle d'interaction spatiale est utilisé pour estimer ou prévoir l'importance des échanges entre couples de lieux*» (*op.cit.*). Il permet donc de mettre en évidence des structures intrinsèques aux échanges qui ne sont pas perceptibles par la cartographie directe des relations observées. Pour y parvenir, il convient de préparer en amont les données en les modélisant à l'aide d'hypothèses et de variables auxiliaires.

Les différents objectifs d'une carte de flux évoqués dans les définitions présentées ci-dessus ont recours à de multiples notions et à plusieurs méthodes de cartographie sur lesquelles nous reviendrons dans la première section du chapitre 4. Aussi allons-nous restreindre notre propos, dans cette section du Chapitre 2, aux seuls aspects graphiques de la carte de flux. L'attention est aussi attirée sur le fait que les aspects graphiques qui portent sur la conception de la carte de flux à savoir l'agencement des figurés et les principes de la sémiologie graphique classique des flux sont présentés dans la seconde section du chapitre 4.

2.3.2 Le principe

 Dresser une carte de flux consiste à projeter leur structure dans un espace métrique, généralement euclidien, muni d'un système de projection cartographique. A l'instar de la représentation sous la forme de graphe, présentée sur la Figure 2-8 (page 68), la cartographie de l'ensemble des valeurs non nulles d'une matrice de flux est généralement illisible, comme illustré par la Figure 2-20.

Figure 2-20. Une carte des flux de l'ex-Tchécoslovaquie en 1990 (version 1)



La Figure 2-20 décrit des liens qui sont plutôt longs, fins et fortement enchevêtrés. Ces liens sont très nombreux, ils semblent bien relier des points précis qu'il n'est pas possible de distinguer.

D'après la légende, ces liens présentent des tailles variables qui sont fonction de la valeur du flux correspondant, mais cette variation de taille n'est pas visible, en raison de la densité visuelle trop importante de figurés. Cette Figure 2-20 est caractéristique du problème de *complexité graphique* déjà mentionné au sujet des graphes (voir section 2.3.1.2, page 80), que nous allons maintenant examiner plus en détail. En effet, cette complexité visuelle constitue la première des contraintes ((carto)graphiques) à résoudre lorsque la matrice est dense (c'est-à-dire composée de plusieurs lieux).

2.3.3 Les contraintes

La représentation cartographique directe d'une matrice de flux dense est caractérisée par une *complexité graphique* intrinsèque liée à la difficulté de représenter (et de percevoir) les changements dans l'espace dont il est question. Les auteurs qui se sont penchés sur la question ont qualifié ce phénomène graphique⁷⁴, d'« absurde » (Bertin, 1973), de « chaos » (Tobler, 1987), de « toile d'araignée des flux » (Dorling, 1991 : 104 ; Dent, 1999 : 105) ou encore « d'effet spaghetti », terme que nous faisons notre pour qualifier la complexité graphique de la figure. Cet effet est défini comme suit par J. Breukelmann et al. « *S'il y a beaucoup de flux, sur une carte, entre les origines et les destinations, leur tracé peut entraîner un motif 'spaghetti' ; tous les types de flux étant dessinés les uns sur les autres, brouillant les motifs généraux. Pour être en mesure de produire une carte lisible, seuls les principaux flux seront par exemple représentés* » (Breukelmann et al., 2009 : 43).

L'effet spaghetti correspond, de notre point de vue, au premier problème de la cartographie des matrices de flux, du moins celui qu'il convient de résoudre en premier. C'est la raison pour laquelle nous allons l'examiner en détail, dans la sous-section suivante. Il convient toutefois de noter que **les possibilités graphiques de la cartographie de flux sont aussi limitées en raison d'une autre contrainte forte liée, quant à elle, à la prise en compte de la composante spatiale des mobilités.** Cette contrainte interdisant formellement le repositionnement des lieux (de manière à minimiser les chevauchements entre les liens), ce qui conduit à renforcer la complexité visuelle de la figure.

De ce fait, **le défi général de la cartographie de flux consiste à réaliser une figure lisible, tout en intégrant d'une part, la contrainte de localisation des lieux que nous examinons dans la seconde sous-section et d'autre part, la richesse de l'information de flux qui porte à la fois sur des lieux et sur des liens**, cette troisième contrainte est examinée dans la dernière sous-section.

2.3.3.1 La complexité graphique ou effet-spaghetti

L'*effet spaghetti* qui caractérise la complexité graphique de la carte de flux issue d'une matrice dense est la manifestation d'un problème beaucoup plus complexe qu'il n'y paraît. Par rapport à l'analyse de la complexité de la carte que D. Fairbain (2006) a proposée, nous appréhendons la complexité graphique de la carte de flux comme une difficulté liée à la compréhension de l'information représentée, en raison du nombre et de la densité de figurés représentés. Dans une carte de flux, la complexité graphique est effective lorsque message apparaît brouillé (incompréhensible, illisible) et non clair (lisible) en raison d'une quantité trop importante de figurés superposés.

De ce point de vue, la résolution intéresse différents aspects de la conception cartographique : ils sont liés au traitement des données numériques, cartographiques et à leur représentation sur une carte. Seul le dernier point nous intéresse à ce stade nous reviendrons sur les aspects liés au traitement des données dans le chapitre 3.

Sur le plan graphique, *l'effet spaghetti* correspond à une densité importante de figurés, par ailleurs fortement enchevêtrés, qui s'avère être la manifestation conjointe de deux effets-graphiques distincts. Le premier d'entre eux, que nous qualifions de **surcharge graphique**, est lié à la quantité de figurés présents sur la carte ; il correspond à un chevauchement ou une superposition trop important(e), souvent renforcé(e) par la présence de longs figurés qui présentant des orientations variés. De fait, lorsque la charge graphique, c'est-à-dire le nombre de figurés, est trop importante, le message est brouillé, illisible.

⁷⁴ Nous verrons, dans la seconde partie, que la cause de ce phénomène de complexité graphique n'est pas exclusivement graphique et que de fait, les solutions qui peuvent être apportées sont d'ordre graphique et sémiologique (Chapitre 5) mais aussi statistiques (chapitres 7 et 8).

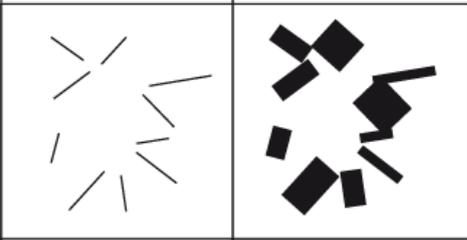
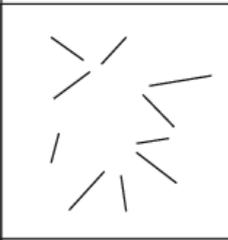
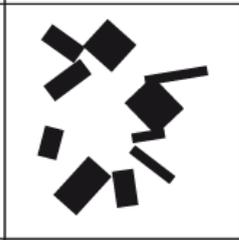
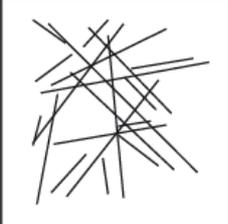
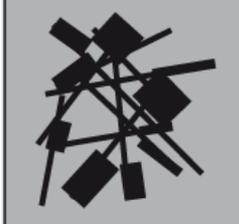
Le second effet, que nous nommons degré de *couverture spatiale*, est lié à la dimension géométrique des figurés défini comme le produit de leur *longueur* par leur *largeur*. Le degré de couverture spatiale caractérise la quantité de noir qui est visible sur la figure, par rapport à celle de blanc. Lorsqu'elle est trop importante, la *couverture spatiale* empêche la perception « claire » des motifs de flux.

Si la surcharge graphique peut être interprétée comme une densité de figurés, la couverture spatiale correspond, quand à elle, à une densité de teinte, ici noire. La Figure 2-21 illustre ces deux effets graphiques dont la combinaison définit *l'effet spaghetti* (ce qui correspond à la partie grisée de la figure).

La densité de figurés correspond au rapport entre le nombre de liens et la part de fond de carte visible. Elle est indépendante de leur taille : la *surcharge graphique* qui en résulte caractérise donc les figurés courts ou bien long. À l'inverse, la densité de noir correspond à un degré de recouvrement du fond de carte par des figurés qui, lorsqu'ils sont trop épais, conduit à assombrir la figure et cela, indépendamment du fait qu'ils soient trop nombreux ou non : la couverture spatiale étant, dans ce cas, d'autant plus importante que les figurés sont larges.

Plus généralement, la *couverture spatiale* des flux dépend de leur nombre et de leur orientation, quand la *surcharge graphique* est liée à leur dimension (à leur *largeur* et à leur *longueur*). Tous ces critères qui sont liés au dessin du figuré décrivant le flux, en l'occurrence le *trait*, varient en fonction du nombre de lieux (*l'effet spaghetti* se manifeste surtout lorsque le nombre de lieux dépasse la dizaine) ; de leur position relative les uns par rapport aux autres (c'est-à-dire du système de projection cartographique) et des caractéristiques spatiales des couples de lieux (de la métrique associée à leur espacement). Des critères perceptifs entrent également en compte puisque l'œil humain ne sait distinguer plus de 7 à 8 variations de forme et de taille. C'est d'ailleurs la raison pour laquelle la figure est considérée comme « *illisible, donc inutile* » (Torricelli, 1997: 9).

Figure 2-21. Présentation de l'effet-spaghetti caractéristique de la cartographie des matrices de flux denses

		Couverture spatiale	
		figurés fins	figurés gros
Surcharge graphique	figurés épars	-	+
			
	figurés nombreux	+	
			

Françoise Bahoken, 2015.

La *complexité graphique* avérée de la cartographie des matrices denses constitue un réel problème, car l'efficacité d'une carte passe par la perception instantanée de l'information qu'elle est supposée transmettre. « *L'excellence graphique correspond à une présentation bien conçue des données d'intérêt [...], [elle] vise à communiquer des idées complexes avec clarté, précision et efficacité ;*

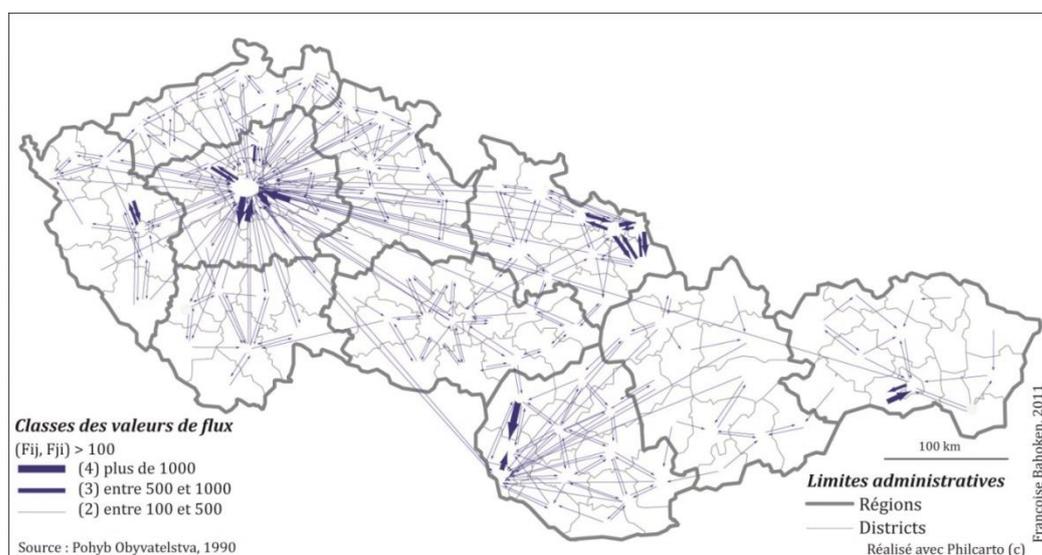
[elle] est ce qui donne au lecteur le plus grand nombre d'idées dans le temps le plus court, tout en utilisant le moins d'encre sur un minimum d'espace » (Tufte, 1983 : 51). Une transmission adéquate du message réside, par conséquent, dans l'articulation d'éléments d'ordres variés qui visent à réduire sa complexité visuelle (Mac Eachren, 1982).

Le raisonnement qui est traditionnellement adopté pour réduire cette complexité graphique consiste à agir au niveau analytique, lors du traitement des données de flux et non particulièrement au niveau graphique, comme nous le montrerons dans le chapitre 5. **La réponse apportée à l'effet spaghetti est donc mise en œuvre, dans l'approche usuelle, en amont de l'étape de représentation cartographique.** Trois grandes voies méthodologiques sont effectivement envisageables.

La première voie consiste à réduire la résolution des données, en utilisant des méthodes de regroupement des objets liées à l'agrégation des données. Ces méthodes réduisent le nombre de lieux (agrégation spatiale) et/ou de liens (typologie par les flux). Etant donné que les méthodes de partition conduisent à un changement conceptuel fort (la cartographie de flux intercommunaux, par exemple, diffère de celle de flux interrégionaux au niveau de l'échelle – locale / régionale – et de l'interprétation – vision locale / nationale), nous n'allons pas développer cet aspect.

La seconde voie est suggérée par W. Bunge (1962, 1969). L'auteur a proposé de résoudre le problème spécifique de la quantité d'information de flux à représenter en la simplifiant, c'est-à-dire en sélectionnant l'information à représenter avec « parcimonie ». Dans ce registre, l'approche la plus courante consiste à éliminer brutalement une partie de l'information, afin de ne représenter que les flux qui sont considérés comme « significatifs ». C'est le cas de l'essentiel des cartes de flux qui ne représentent que les n flux les plus importants (en valeur, en tonnage, en volume...) ou qui sont supérieurs à un seuil donné. La carte de la Figure 2-22 illustre ce principe de cartographie d'une sélection de valeurs de flux supérieures à 100 pour les flux de l'ex-Tchécoslovaquie (1990). Elle correspond à une simplification de l'effet spaghetti présenté sur la carte de la Figure 2-20 (page 85).

Figure 2-22. Cartographie des valeurs de flux supérieures à 100 migrants, dans l'ex-Tchécoslovaquie, de 1990



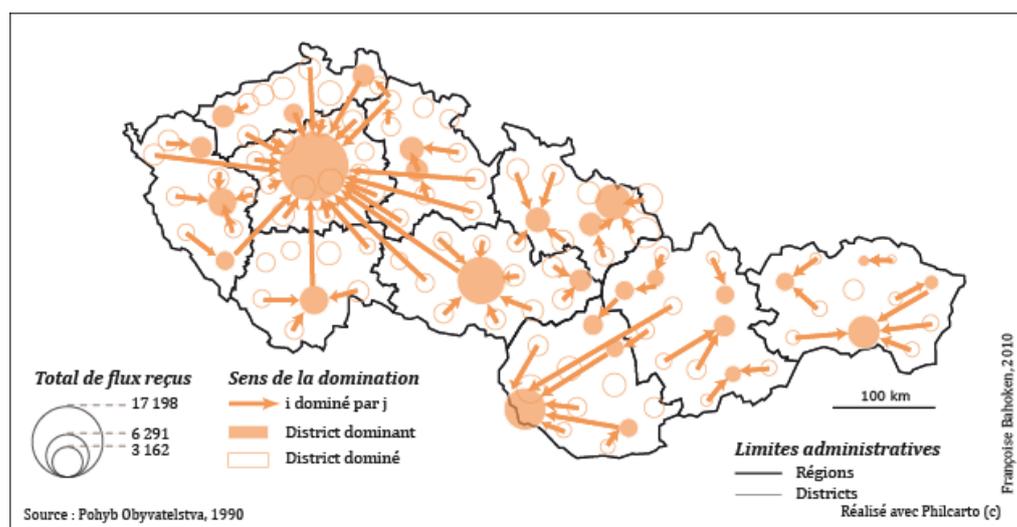
La question de la quantité de flux à représenter sur une carte faisant appel à celle du seuillage de données, elle n'est pas aisée à résoudre au premier abord, en raison des nombreux paramètres qui entrent en jeu. C'est la raison pour laquelle nous y reviendrons dans les chapitres 7, 8 et 9 de la seconde partie.

La troisième voie méthodologique conduisant à réduire la complexité graphique de la carte de flux est suggérée par M. Beguin et D. Pumain (1994). Elle s'appuie sur des considérations qui sont davantage analytiques que graphiques. « *Les représentations cartographiques directes de tableaux d'échanges deviennent difficiles à construire et à lire lorsque le nombre des unités géographiques considérées est grand. Dans ce cas, on a intérêt à traiter l'information statistique avant de la représenter (par exemple en identifiant les flux principaux, ou en ajustant le tableau d'échanges par un modèle gravitaire pour faire apparaître par différence des flux préférentiels ou des effets de barrière)* » (Beguin, Pumain, 1994 : 118).

Les auteures proposent, à titre d'exemples, deux approches méthodologiques qui sont s'inscrivent dans les cadres d'analyse descriptive et explicative des flux. L'approche descriptive consiste à sélectionner les flux de manière à ne représenter que ceux qui sont « *principaux* », ou bien ceux que l'on considère comme *significatifs* ou *représentatifs* du système d'échanges ; qui sont *majeurs* ou *dominants*, tout dépend du critère retenu.

La Figure 2-23 propose ainsi une représentation des *flux dominants* de l'ex-Tchécoslovaquie (1990). L'attention est attirée sur le fait qu'il s'agit là d'une interprétation cartographique de la notion de *flux dominants* proposée par D. Nuysten et M. F. Dacey (1961), nous y reviendrons dans le chapitre 9.

Figure 2-23. Cartographie des flux dominants des migrations de l'ex-Tchécoslovaquie, en 1990



La seconde approche, davantage explicative, a pour objet de ne représenter que les *flux préférentiels* (à l'inverse, des *effets de barrières*) ou, plutôt, une *sélection significative d'entre eux*, ces derniers ne pouvant, dans bien des cas, être perceptibles sans sélection préalable.

Ces deux dernières voies méthodologiques caractérisent une démarche fondée sur la sélection préalable des valeurs de flux avant leur cartographie, que celles-ci aient fait l'objet – ou non – d'un traitement particulier (visant, par exemple, à extraire les flux préférentiels de la matrice des données initiales). C'est pourquoi le principe général de réalisation d'une carte de flux consiste à sélectionner au préalable les données à représenter, afin d'éviter *l'effet spaghetti*. Son efficacité n'est pas à démontrer, car l'on voit bien sur la Figure 2-23 l'effet de la seule représentation des *flux dominants* sur l'efficacité visuelle et cognitive de la figure : nous reviendrons sur cette méthode dans le chapitre 9. Sur le plan thématique, cette carte met bien en évidence la structure multipolaire organisée, au niveau supérieur, autour de trois grands pôles principaux qui correspondent aux anciennes capitales historiques ; ces derniers sont relayés au niveau local par quelques pôles régionaux.

L'application d'un critère de sélection des valeurs de flux à cartographier est une pratique courante en cartographie de flux. Cependant, la procédure ne concerne pas le seul filtrage des valeurs de flux comme nous venons de le voir, elle intervient à différents étapes de la production cartographique.

2.3.3.2 La sélection des données de flux

La sélection des données de flux peut être appréhendée comme une réflexion sur la *focale* de la matrice, c'est-à-dire la (ou les) partie(s) de la matrice qui est (sont) concerné(es) par la cartographie, éventuellement de la manière dont l'espace est pris en compte. Dans ce cas, la sélection porte sur :

- les lieux d'origine et/ou de destination (O_i , D_j) ;
- les liaisons entre les lieux (L_{ij}) ;
- la valeur des flux entre les lieux (F_{ij}), ou celle de leur intensité lorsque les données sont transformées en indicateur ;
- l'espace des flux, appréhendé par exemple, en termes de termes de *distance parcourue par le flux* (D_{ij}) ou encore de coût de déplacement (K_{ij}) ;
- un raisonnement mixte qui prend en compte flux et distance parcourue (F_{ij} , D_{ij}) peut également être mis en œuvre.

Lorsque les flux sont appréhendés du point de vue des seuls lieux, la carte vise à mettre en évidence l'effet des flux sur la distribution des lieux, elle correspond alors à une variante que nous présenterons dans la section 2.3.4.1.

La sélection des objets peut aussi être réalisée lors du traitement des données cartographiques. Le choix porte alors, par exemple, sur le type de projection du fond de carte, éventuellement sur sa métrique (euclidienne ou non) ; sur la modélisation des objets géographiques et de leurs relations. Dans notre cas, nous faisons le choix *a priori* d'utiliser le trait droit pour représenter le flux et le point, pour illustrer le lieu. Nous reviendrons sur les différentes possibilités graphiques et esthétiques dans le chapitre 4.

Ce principe de sélection des données de flux au préalable de leur représentation n'est pas propre à la représentation cartographique : nous avons vu qu'il était également valable pour le graphe ordinaire. Les méthodes mises en œuvre pour réduire la complexité visuelle de la carte de flux sont donc en partie similaires à celles du graphe (il s'agit en particulier des solutions graphiques, d'ordre visuel et/ou liées à l'esthétique de la figure que nous avons présenté à la section 2.2), à l'exception de celles qui conduisent au repositionnement des lieux (des sommets) de manière à réduire le chevauchement des liens. En effet, ces solutions typiques de la représentation des matrices sous la forme de graphe, ne sont pas applicables à la cartographie de flux : le changement de la position des lieux sur une carte ne pouvant être ni aléatoire ni arbitraire, il est nécessairement associé à une transformation cartographique de position, à un changement de projection qui conduit à une redistribution relative des positions.

La composante spatiale constitue par conséquent la seconde contrainte forte à la réalisation d'une carte de flux. Cette seconde contrainte intéresse les questions liées à la position des lieux dans l'espace que nous examinons dans la sous-section suivante. Cette position des lieux contraint, *a fortiori*, la représentation des flux selon le modèle *lien-nœud* qui consiste à combiner, sur la même figure, la représentation des flux observés du point de vue des liens et des lieux.

2.3.3.3 La localisation des lieux



La contrainte forte de localisation des lieux est liée à la projection, dans l'espace géographique, de la structure de flux ancré sur des nœuds – des lieux – fixes et entre lesquels

sont placés des arêtes – ou arcs – qui en deviennent, elles-aussi fixes. On observe une certaine rigidité dans la construction d'une carte de flux liée à la structure topologique⁷⁵ de l'espace lorsqu'il est appréhendé de manière discrète.

L'espace dont il est question dans une carte est toujours représenté par un fond de carte, à une échelle donnée. Ce fond est composé d'objets géographiques dont la position est absolue dans un système de coordonnées précises. Les possibilités de déplacement des lieux, hormis le cas d'un changement de référentiel cartographique ou d'une transformation⁷⁶ du fond de carte sont, par conséquent, limitées. Les possibilités de placement des liens entre des couples de lieux sont, de fait, réduites par des positions des lieux déterminées *a priori*.

Certaines contraintes esthétiques applicables aux graphes sont ainsi hors de propos pour les cartes de flux : c'est le cas du placement optimal des nœuds ou des liens, pour assurer la lisibilité de la figure, ou encore de la minimisation des chevauchements autrement que par l'application d'angles au tracé des arcs. Ceci étant, il est possible de modifier la position des lieux d'un espace d'étude sous réserve, d'abandonner, par exemple, le cadre de la géométrie euclidienne classique sans nuire à la lisibilité de l'information. L'application CARTOD (Bonin, Bahoken, Mermet, 2013), en opérant une transformation cartographique de la position des lieux d'origine et de destination des flux permet justement de réduire significativement la complexité graphique de la carte. Ces contraintes liées à la composante spatiale de localisation sont spécifiques à la carte de flux et secondaires à la réalisation d'un graphe des mêmes flux (la théorie des graphes faisant abstraction de l'espace). Il est ainsi toujours possible de représenter graphiquement l'ensemble des liaisons entre les N sommets à l'aide d'une matrice ordonnée ou d'un graphe, en appliquant pour ce dernier des contraintes esthétiques et/ou de déplacement des nœuds par une force qui agit sur leurs positions, ce qui n'est pas le cas en cartographie des flux, la position des lieux ne pouvant être modifiée *a priori*.

Dans un graphe non planaire, la position des nœuds n'a pas toujours de sens alors qu'elle est fondamentale dans une carte de flux. C'est la raison pour laquelle nous l'avons introduite comme un cas particulier de graphe. Cette différence liée à la topologie des lieux présente des conséquences sur la figure toute entière : les algorithmes de visualisation des graphes permettent l'optimisation du placement des lieux sur le plan esthétique, pour privilégier l'affichage des liens existants, au détriment de la position des lieux, quand les outils de cartographie de flux se distinguent par la manière dont ils sélectionnent et représentent les liens, en fonction de leur valeur (Tobler, 1987 ; Dorling, 1991, Glennon et Goodchild, 2004 ; Van der Zwan, 2005).

La question de savoir si les liens existent ou non ne se pose pas en ces termes en cartographie de flux. Tout d'abord, en fonction de l'échelon d'observation des flux, il est tout à fait possible d'avoir à représenter tous les liens qui peuvent exister entre les unités spatiales d'une zone d'étude : tout dépend de la manière dont sont appréhendés les flux à représenter, du nombre de flux à représenter et de l'objectif de la carte : *S'agit-il de représenter des flux du point de vue des lieux d'origine (ou de destination) ou bien du point de vue des relations entre les lieux ?*. Cette question est liée au choix d'une focale d'analyse de la matrice de flux.

Ensuite, dans une carte de flux, il importe moins de représenter l'existence ou l'absence de relation entre des lieux que la valeur de leurs relations, c'est-à-dire la quantité (Fij) qui s'est déplacée : un effectif de migrants, un nombre de migrations, un volume, un tonnage ou encore une valeur (en euros, par exemple). Il s'agit précisément de rendre compte graphiquement des différences de

⁷⁵ La topologie est une branche des mathématiques qui fournit un schéma conceptuel invariant permettant de décrire, dans un espace métrique, les relations entre des objets telles que la connectivité.

⁷⁶ Il s'agit ici de la famille des transformations cartographiques de position (Cauvin, Enaux, 1991 ; Cauvin, 1998 ; Denain, Langlois, 1998) qui constitue l'une des variantes envisageables de la carte des flux.

proportionnalité entre ces différentes relations spatialisées, en utilisant *a minima* une variation de la taille des figurés – des liens et des points – qui les symbolisent, par le recours à la variable visuelle *Taille* telle qu'elle est définie par J. Bertin (1967). Cet objectif est rendu difficile à atteindre, lorsque la matrice de flux est dense. C'est notamment la raison pour laquelle la question de l'existence des relations présentes sur la figure résulte généralement de l'application d'une méthode ayant pour objectif de limiter le nombre de liens ou de flux à représenter.

Cette question de la quantité de flux à représenter, que nous avons déjà observée au niveau des liens (voir Figure 2-21) se manifeste avec encore plus d'acuité lorsqu'elle est combinée sur la même carte avec celle de la symbolisation des lieux.

2.3.3.4 La combinaison des liens et des nœuds

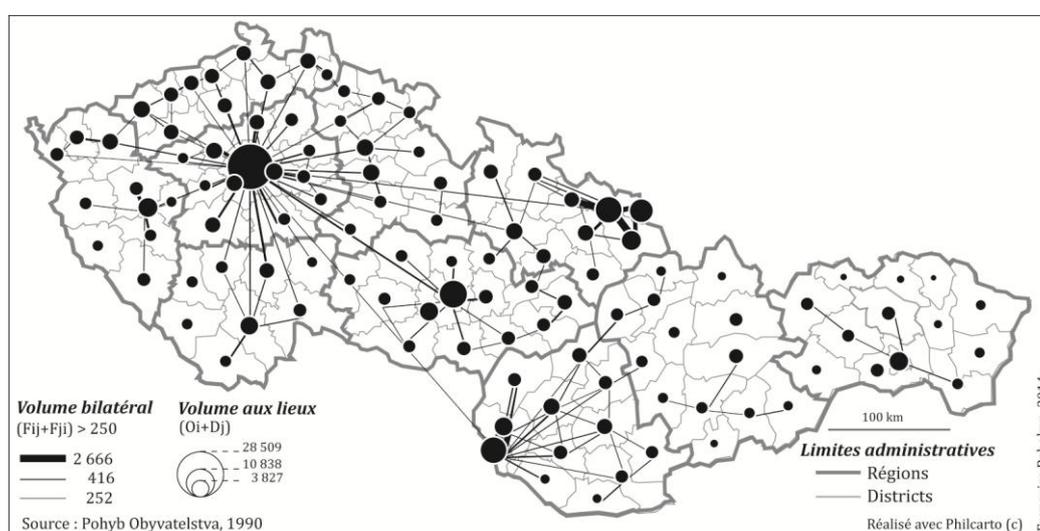


Sur le plan graphique, la pratique qui consiste à ne représenter que les n flux les plus importants qui se sont produits entre des couples de lieux est justifiée par des impératifs graphiques de placement des différents figurés, mais également par la nécessité de combiner, sur une même figure, un premier groupe de symboles qui illustrent les relations entre les lieux et un second groupe qui représente l'effet de ces relations sur la distribution des lieux. En d'autres termes, cela consiste à représenter, à l'instar du graphe non planaire, des symboles proportionnels (ou pas) à un indicateur de l'effet des flux aux lieux (le volume, par exemple) et de les placer sur les lieux.

La combinaison de la représentation des liens et des lieux sur la carte de flux engendre une difficulté graphique supplémentaire puisque doivent cohabiter deux types de symboles différents sur la même figure : l'un pour les lieux et l'autre pour leurs relations. Cette contrainte est importante à intégrer, car l'analyse spatiale des flux réalisée en amont de la représentation vise généralement à mettre en évidence, soit des processus de différenciation spatiale, dans le cadre d'un raisonnement mené au niveau des lieux, sur l'effet des flux sur la distribution des lieux, soit d'analyser les interactions spatiales dans le cadre d'un raisonnement mené sur les valeurs ou sur l'intensité des flux.

Combiner la représentation des liens et des lieux consiste à donner à voir la même information, sous deux angles différents : du point de vue des liens et de celui des lieux. Cela signifie aussi de représenter le même processus, sous deux formes distinctes mais complémentaires. Si celles-ci représentent effectivement le même phénomène, les deux informations qui portent sur les liens et lieux ne sont pas pour autant redondantes (voir Figure 2-24).

Figure 2-24. Cartographie du volume bilatéral de migrations de l'ex-Tchécoslovaquie en 1990



La Figure 2-24 illustre le volume des flux bilatéral (du point de vue des liens) combiné au volume de flux aux lieux, de l'ex-Tchécoslovaquie (1990). Les symboles ponctuels placés sur les lieux sont proportionnels au volume de flux aux lieux ($V_i = O_i + D_j$), tout comme les liens placés entre les lieux auront une épaisseur proportionnelle au volume bilatéral des flux ($V = F_{ij} + F_{ji}$).

La carte de flux que nous venons de décrire correspond au cas général. Elle est typique du modèle général que J.-C. Müller (1991) qualifie de *modèle des atlas*. Il en existe des variantes, souvent peu communes dans les atlas évoqués, qui diffèrent du cas général par les choix réalisés par l'auteur, au moment de l'étape de représentation.

Face aux difficultés d'ordre méthodologique liées à la représentation des liens, différents auteurs ont proposé la réalisation de variantes à la carte de flux classique. Il en existe de nombreuses, souvent peu communes, et absente des atlas évoqués. Ces variantes diffèrent du cas général par les choix réalisés par l'auteur, au moment de l'étape de représentation. Leurs fondements méthodologiques s'appuient sur des aspects méthodologiques qui sont liés au contexte cartographique. Ces choix sont effectivement susceptibles de faire varier la figure, car ils portent sur la partie de la matrice à représenter, sur la représentation de l'espace qui n'est pas nécessairement conventionnelle et sur celle des liens.

2.3.4 Les variantes

Les choix méthodologiques réalisés par l'auteur d'une carte statistique sont susceptibles de faire varier la figure. Nous les avons regroupés en trois catégories assez générales qui correspondent à une action menée : i) sur la partie de la matrice à représenter ; ii) sur la manière de représenter l'espace, qui n'est pas nécessairement conventionnelle et iii) sur la façon de tracer les liens qui illustrent les flux. Ces trois catégories correspondent à autant de manières d'aborder l'objectif de cartographie d'une matrice de flux.

Si nous avons présenté la carte de flux (liens, lieux) comme le modèle général, il est important de garder à l'esprit que les cartes de flux disponibles correspondent souvent à la représentation de l'effet des flux sur la distribution des seuls lieux (d'origine et/ou de destination). Dans ce cas, l'auteur a fait le choix de ne sélectionner qu'un seul type d'objet géographique, les lieux et non les liens, pour représenter les flux dont il est question. Cela signifie que les liens seront absents de la représentation graphique, leur valeur pouvant toutefois être intégrée sous la forme d'indicateur. Cette cartographie des flux focalisée sur les lieux est aussi susceptible de varier en fonction du type de représentation cartographique.

Un même phénomène localisé peut ainsi faire l'objet de plusieurs variantes qui dépendent de la manière dont la carte est réalisée. De ce point de vue, *l'Atlas Interactif des Régions Européennes (AIRE)*⁷⁷, développé par l'UMS 2414 RIATE (Zanin, Lambert, 2009), est un bon exemple de la variation de la figure liée aux choix conceptuels et méthodologiques des auteurs. En effet, la figure varie dans cet atlas en fonction du type de données (de stocks ou de rapports) ; du type de fond de carte caractéristique d'une implantation spatiale des lieux (sous la forme de points ou de surfaces) et d'une éventuelle normalisation. Les fonds de carte d'origine administrative, réalisés en implantation surfacique, sont souvent substitués à d'autres formes de représentation de l'espace. Ils sont soit remplacés par des maillages réguliers qui prennent diverses formes (carrées, rectangulaires ou encore en nids d'abeille), soit transformés par l'application d'une procédure d'anamorphose scalaire ou d'une

⁷⁷ Cette application est disponible à l'adresse suivante : <http://aire.ums-riate.fr/fr#content/index.fr.php> [lien érigé le 02/08/2015].

interpolation de manière à ce que les données puissent être représentées de manière continue, indépendamment des limites administratives.

Les choix qui portent sur la manière de représenter l'espace correspondent à sa formalisation graphique : sous la forme d'un plan, en deux dimensions, ou bien d'une ligne : c'est-à-dire en une dimension. Enfin, les choix qui concernent la représentation des liens portent sur les aspects graphiques liés au tracé. Le tracé du lien est généralement droit ou courbe, ce qui revient à première vue au même. Les aspects graphiques dont il est question ici concerne les algorithmes de fusion graphique des liens, soit parce qu'ils ont été sélectionnés, soit parce qu'ils sont réorientés en fonction de leurs valeurs.

Ces différents choix, réalisés par l'auteur au moment de la réalisation de la carte, ne sont pas anodins : ils affectent la figure et par conséquent les motifs représentés, leur perception, sans toutefois être toujours efficaces. Nous présentons tour à tour des exemples des trois familles de variantes.

2.3.4.1 *Le point de vue des lieux*

La première famille de variante de la carte de flux réalisée selon un modèle lien-nœud vise à représenter les échanges du point de vue des seuls lieux. Cette opération consiste à un changement de focale d'observation des échanges puisque ce n'est plus tant la relation entre les lieux qui est représentée – la valeur de la cellule (i,j) – que le résultat de cette relation au niveau des lieux d'origine (i) ou de destination (j). Le changement de focale résulte donc d'un choix fort, celui « [...] *de ne cartographier que les flux qui ont une même unité géographique pour origine (ou pour destination). On se ramène alors à une implantation ponctuelle de l'information, et, plutôt que de figurer les flux par des flèches, il est possible de recourir à un figuré ponctuel : par exemple, on représente par des cercles de surface proportionnelle le nombre de migrants vers Paris originaire de chaque département* » (Beguin, Pumain, 1994 : 118). Dans l'exemple d'application proposé par M. Beguin et D. Pumain, la carte est alors composée de points proportionnels à l'effectif de migrants installés à Paris. Les symboles, ponctuels et non plus linéaires, sont placés sur chacun des départements, plus précisément au centre géométrique des unités spatiales qui les représentent.

La réalisation d'une telle carte représente le modèle couramment utilisé. Il faut reconnaître qu'elle présente de nombreux avantages. Tout d'abord, elle résulte du choix de l'auteur de focaliser l'analyse sur les seuls flux qui présentent la même origine ou destination, ce qui permet de raisonner du point de vue des lieux et non plus de leurs relations. Elle présente aussi l'intérêt de produire une image lisible, même si elle ne concerne qu'une partie de l'information disponible : l'information sur les échanges entre les lieux est, de fait, perdue au profit de celle sur des lieux. Le gain apporté réside dans l'annulation du problème de *l'effet-spaghetti*, ce qui justifie l'abondance de telles cartes.

Le procédé permet enfin de produire, à peu de frais, une représentation des échanges qui ne nécessite pas la manipulation de matrices de flux souvent considérée comme rédhibitoire. Sa réalisation devient aisée : les méthodes de la cartographie statistique sont disponibles dans plusieurs ouvrages (Beguin, Pumain, 1994, 2003 ; Cauvin, Escobar, Serradj, 2008, *tomes 1 et 3*, Le Fur, 2007 ; Le Fur et al., 2010 ; Joly, 1976 ; Pumain, Saint-Julien, 2001 ; Weger, 1999 ; Zanin, Trémélo, 2003) et les modes de représentation variés, comme l'illustre la Figure 2-25.

Figure 2-25. Les procédés de représentation des flux, du point de vue des lieux

Etat et variations des échanges aux lieux								
symbole proportionnel (1)	symbole proportionnel (2)	semis de points	Plus-moins	aplats	tendances	gradients isolignes	gradients potentiel	volume proportionnel

©) Françoise BAHOKEN, 2010

Pour autant, lorsque l'on combine la représentation de symboles proportionnels aux liens et aux lieux, la carte apporte plusieurs informations complémentaires, comme nous l'avons vu dans la section précédente.



W. Tobler a proposé un ensemble de variantes à la méthode des symboles proportionnels (Tobler, 1979a, 1981, 1995, 2001a). L'une d'entre elles consiste à représenter par des signes positifs (+) et négatifs (-), la polarité du résultat du calcul d'un indicateur des flux, la migration nette, par exemple, calculé aux lieux. Ces symboles, proportionnels à la valeur de l'indicateur, sont placés au centre des zones. La Figure 2-26 illustre l'application de procédé à la cartographie du solde des flux aux lieux de l'ex-Tchécoslovaquie (1990). L'auteur justifie cette proposition sémiologique en montrant tout l'intérêt de son utilisation dans le cadre d'observation de la répartition spatiale des fréquences d'occurrence des différents signes. L'inversion de polarité entre des lieux contigus (Tobler, 1995 : 330) permettant de révéler, notamment, l'existence d'une auto-corrélation spatiale.

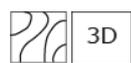
Figure 2-26. Une carte des flux de l'ex-Tchécoslovaquie de 1990, en symboles plus-moins



D'autres variantes proposées par l'auteur méritent également l'attention. C'est le cas, par exemple, des représentations en trois dimensions qui consistent à extruder la zone de manière proportionnelle à la valeur du stock. Dans les versions proposées par l'auteur, la transformation est souvent associée à une transformation de l'espace, à son acception en tant que surface continue et non plus discrète.

Ces deux méthodes de transformations des données cartographiques sont présentées dans la sous-section suivante.

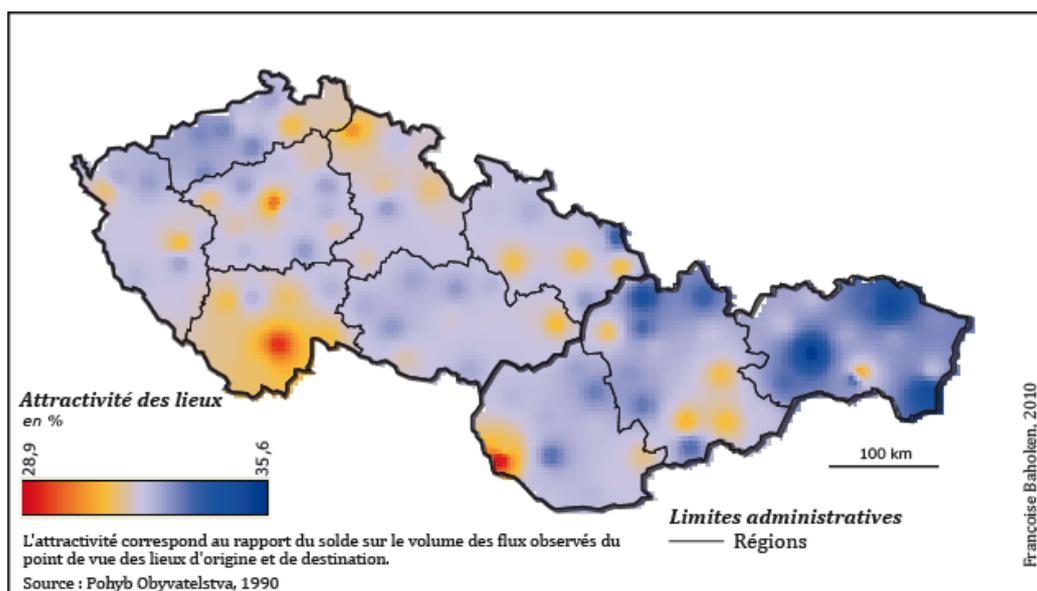
2.3.4.2 La transformation d'implantation spatiale et l'interpolation



W. Tobler (1979b) a proposé la réalisation de cartes de flux selon la méthode qu'il qualifie de « *Pycnophylactic* » (dans le texte). Il s'agit de la représentation interpolée et en trois dimensions d'une carte choroplète. La méthode articulant en fait deux méthodes sophistiquées de représentation : une interpolation et une extrusion en trois dimensions.

La Figure 2-27 présente une cartographie interpolée, en deux dimensions, des flux de l'ex-Tchécoslovaquie exprimant l'attractivité des districts.

Figure 2-27. Une carte des migrations tchécoslovaques (1990) interpolée



La seconde partie de la méthode est une redistribution spatiale des stocks couplée à une extrusion des valeurs. Elle revient à faire une carte en trois dimensions supprimant les pics tout en conservant le volume de chaque zone. La Figure 2-28 illustre la méthode d'interpolation « *Pycnophylactic* » de W. Tobler (1979b).

Figure 2-28. La figure de la carte de type *Pycnophylactic* (Tobler, 1979)



Source : W. Tobler, 1979.

Outre le changement d'implantation, la transformation de l'espace peut également être appréhendée comme un changement de représentation.

2.3.4.3 La transformation de l'espace par l'anamorphose

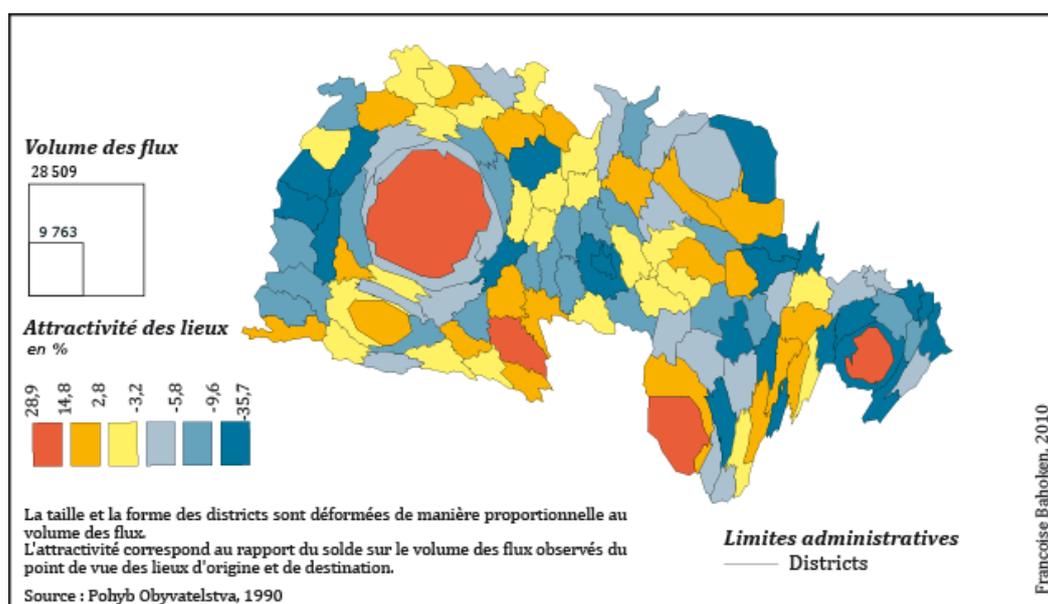
Une première manière de transformer l'espace consiste à appliquer une procédure d'anamorphose, la seconde possibilité consistant à représenter l'espace sous la forme d'une ligne.

La procédure d'anamorphose, appliquée aux flux, a été utilisée pour représenter le mouvement de manière dynamique sur un support statique. C'est pourquoi les cartes ainsi créées ont été qualifiées de « *stato-dynamiques* », par S. Rimbart (1990).

L'idée de représenter la dynamique du temps par un procédé d'anamorphose consiste, d'après G. Palsky (2004a), à transposer l'idée de diagramme à la cartographie. Pourrait-il en être autrement car, comme disait Scutenaire cité par l'auteur : « [...] avec les mots on marque le mouvement, avec les images on les fixe » (Palsky, 2004a : 347). Le procédé de l'anamorphose donne ainsi un statut intermédiaire aux cartes censées représenter un mouvement dans l'espace.

L'anamorphose cartographique⁷⁸ consiste, dans sa version scalaire, à déformer le fond de carte en fonction d'un indicateur. La Figure 2-29 représente l'attractivité des districts de l'ex-Tchécoslovaquie, sous la forme d'une anamorphose scalaire, elle constitue de ce fait une variante de la Figure 2-27. Cette anamorphose scalaire est réalisée du point de vue de la masse des lieux. Elle diffère de l'anamorphose vectorielle qui déforme l'espacement entre les lieux en fonction d'une acception de leur accessibilité.

Figure 2-29. Une carte de l'attractivité des flux de l'ex-Tchécoslovaquie de 1990, en anamorphose scalaire



Cette méthode développée à la suite des travaux de T. Hägerstrand (1957), « [...] consiste à transposer, à transcrire une variable descriptive des lieux en une variable déformant ces lieux, en ce sens c'est une métaphore spatiale » (Denain, Langlois, 1997). Elle vise ainsi à utiliser le fond de carte pour représenter le phénomène, ce qui conduit à changer son statut : il n'est dès lors plus utilisé en tant que simple support de représentation du phénomène puisqu'il est le phénomène.

⁷⁸ Pour une analyse comparative des différentes méthodes d'anamorphoses qui peuvent être mises en œuvre pour la réalisation de cartes thématiques, voir notamment : Langlois, P., 2003, Anamorphoses par transformations pseudo-équivalentes, *Actes des VIèmes Rencontres de ThéoQuant*, 17 p. et la note de W.R. Tobler : Pseudo-cartograms.

Il s'agit d'une « *transformation cartographique de position* [...] de type] *morpho-thématique* » (Cauvin, Enaux, 1991) qui s'apparente à un changement de projection de la carte, avec tous les choix que cela implique. Le choix de la projection est ici focalisé sur la réalité des données – et plus sur la position relative des objets les uns par rapport aux autres – afin de permettre une marginalisation des zones où il ne se passe pas grand-chose.

La position relative des lieux les uns par rapport aux autres, mis en interaction spatiale sur une carte de flux, est traduite et estimée dans des matrices spécifiques de position, qui renseignent sur les distances spatiales entre les lieux (Tobler, 1963, 1977). Ces matrices peuvent ensuite être utilisées pour représenter ou non l'espace des lieux. Deux approches sont envisageables d'après W. Bunge (1962) : la première consiste à « [...] *représenter des distances compliquées sur une carte simple* », comme dans l'approche usuelle et la seconde, « [...] *des distances simples sur une carte compliquée* », par le biais d'anamorphoses où la géographie des lieux, exprimée en distance-temps de déplacement ou de transport et non plus selon une métrique euclidienne, est déformée. Cette seconde série de cartes, mise en œuvre par W. Tobler (1963, 1977), C. Cauvin (1996) et A. L'Hostis (1996), s'oppose ainsi à la vision euclidienne classique, en modifiant les positions relatives des lieux et leur représentation graphique. Comme le souligne R. Brunet, « *Certaines cartes deviennent des épures, où les lieux changent de place ou s'éloignent du centre choisi : ces mouvements apparents des lieux sont forts intéressants en eux-mêmes* » (Brunet, 1987 : 100).

La proposition de A. L'Hostis, sous la forme de « *cartes chiffonnées* » a ceci de particulier qu'elle met en évidence un « *relief de distance-temps* » de déplacement en fonction de modes de transport. L'auteur représente ainsi une carte prismatique où les liaisons les plus rapides entre deux villes (liées à l'existence de lignes à grande vitesse routière ou ferroviaire), correspondent aux distances les plus courtes ; elles sont placées au dessus de la figure et relient les crêtes des monts et montagnes correspondant aux villes concernées. A l'inverse, les liens correspondant aux temps de transport les plus longs sont placés en soubassement des monts, ils suivent plus ou moins les fonds des vallées.

Deux types d'anamorphose vectorielle peuvent être réalisés, selon que la distance-temps est calculée depuis un lieu particulier (anamorphose vectorielle unipolaire), ou qu'elle concerne l'ensemble des lieux (anamorphoses vectorielles multipolaires). Si le premier type est souvent réalisé pour illustrer des cartes d'attractivité, liés à l'accessibilité selon différents modes de transport ; celles qui sont spécifiques aux flux sont moins courantes. Les illustrations notamment disponibles dans les travaux de Denain et Langlois (1997), de A. L'Hostis (1997) et de P. Langlois (2006) mobilisent en réalité des *matrices d'adjacence* et non de flux, en référence à des réseaux de transport.

Si le modèle de l'anamorphose considère que le fond de carte déformé se suffit à lui-même pour représenter le phénomène, D. Dorling (1991) a proposé une combinaison intéressante – non reprise à notre connaissance – qui consiste à cartographier les relations entre des lieux sur un fond de carte en anamorphose scalaire.

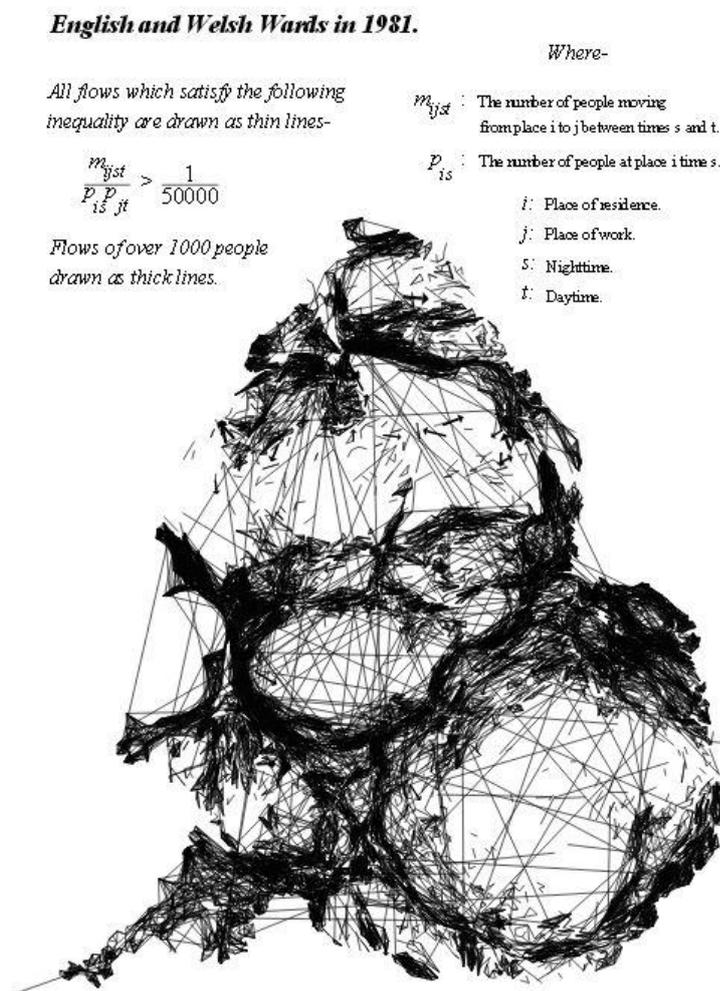
- *Les flux sur fond d'anamorphose scalaire*

L'idée d'utiliser un fond de carte anamorphosé pour représenter des flux s'inscrit dans la recherche de solutions permettant de résoudre les problèmes graphiques liés à la cartographie des flux (Voir chapitre 3). D. Dorling (1991), propose en particulier de représenter les flux sur une anamorphose scalaire où les unités spatiales sont déformées proportionnellement à l'effectif de population des zones.

L'auteur justifie sa proposition en soulignant les propriétés de la représentation en anamorphose pour représenter la réalité du phénomène mais aussi pour donner à voir la vraie densité des lignes qui

illustrent les mouvements (Dorling, 1991 : 113). D. Dorling propose notamment une carte des navettes quotidiennes entre les quartiers britanniques et gallois en 1981, qui est présentée sur la Figure 2-30.

Figure 2-30. Illustration des flux sur un fond d'anamorphose scalaire (Dorling, 1991)



Source : D. Dorling, 1991.

Sur la Figure 2-30, les flux qui correspondent à plus de 2% de la population active des zones de résidence sont représentés par des liens valués, à destination des lieux de travail. Le fond de carte est anamorphosé selon le principe des densités égales où la taille des unités spatiales est proportionnelle à la population active des zones de résidence.

La carte de flux réalisée en anamorphose consiste donc à redistribuer leur structure spatiale dans un espace dont la métrique est fonctionnelle et non plus géographique, associée à un système de projection. Elle modifie les proximités géographiques entre les lieux sur une base cognitive et non plus spatiale. Elle illustre, en ce sens, une autre perception de la réalité des flux car ce qui est cartographié est davantage une structure perçue des relations entre lieux que leur organisation géographique effective. « Elle a pour but de faire apparaître des relations sous-jacentes, d'extraire des caractéristiques non visibles directement » (Cauvin, 1998). C'est notamment le cas de la série de cartes des flux sur fond d'anamorphose proposées par D. Dorling (1991) qui font émerger la structure spatiale particulière des navettes.

Cependant, l'inconvénient du procédé de l'anamorphose cartographique est de n'être pas unique et ce, pour deux raisons principales. La première tient au fait qu'il s'agit de *constructions non reproductibles* (Cauvin, Eaux, 1991) à l'instar des cartes choroplèthes ou issues de procédés similaires de transformations telles que les cartes piézo-plèthes (Cauvin et al., 1995). La seconde raison est liée aux possibilités d'erreurs d'interprétation issues d'une analyse des anamorphoses qui est essentiellement visuelle (MC Carty, Salisbury, 1961). Les anamorphoses réalisées à partir des unités spatiales d'un même territoire ne sont effectivement comparables⁷⁹ que visuellement : elles ne sont pas « [...] comparables d'une manière généralisable, qui conduise à des résultats identiques, quelle que soit la personne qui effectue la comparaison » (Cauvin, Eaux, 1991).

Le modèle de représentation des flux à l'aide des procédés d'anamorphose vectorielle engendre une amélioration nette de la perception d'ensemble des flux par les formes des relations ainsi révélées⁸⁰. Elle autorise une mise en lumière de la dynamique du processus et constitue, à notre sens, un modèle d'illustration des possibilités de représentation dynamique des mouvements, sur un support statique. À l'inverse, la cartographie des flux sur fond d'anamorphose n'est pas toujours simple à analyser (Dorling, 1991 : 114), lorsque le fond est non euclidien.

2.3.4.4 La transformation de l'espace, par le changement de représentation

 La seconde variante à la carte de flux classique, probablement la moins répandue, consiste à ne pas représenter l'espace en deux dimensions, mais sur une seule dimension. Si les économistes peuvent réaliser des graphiques où l'espace est représenté sur une dimension, par une ligne horizontale⁸¹, la représentation dont il est question ici prend la forme d'une ligne verticale. L'espace des lieux d'origine est alors dissocié de celui des lieux de destination, placés tous deux sur des axes dits en coordonnées parallèles.

- *Les flux sur coordonnées parallèles*

Le principe du mode de représentation en coordonnées parallèles (Inselberg, 1985) consiste à positionner sur des axes (thématiques) verticaux parallèles, des nœuds qui illustrent la valeur d'une caractéristique sémantique et à relier les nœuds de différents axes par des lignes droites ou brisées, en fonction du nombre d'axes de coordonnées. Le nombre d'axes parallèles dépend du nombre de variables représentées. Ce procédé est typiquement utilisé pour représenter et comparer les caractéristiques d'un grand nombre d'objets, autorisant ainsi une lecture à la fois globale et locale des données qui sont par nature multidimensionnelles⁸². Il est aussi mis en œuvre dans le cadre d'analyse

⁷⁹ C'est par la méthode de régression bidimensionnelle appliquée à la cartographie, telle qu'elle est proposée par W. Tobler (1986), que les anamorphoses peuvent être comparées.

⁸⁰ Sur ce sujet précis, il convient de noter les travaux de C. Cauvin à l'Université de Strasbourg (Cauvin C., 1997 et 1984) notamment fondés sur l'hypothèse selon laquelle la « [...] forme exprime la structure d'un phénomène ; si l'on est à même de caractériser une forme ou de comparer des formes, alors on peut révéler par l'intermédiaire de la forme une structure non visible » (Cauvin, 1997 : 2). De ce fait, les « transformations [géométriques] permettent de trouver les positions relatives de lieux connaissant des liens entre lieux, qu'il s'agisse de distances-temps, de coûts, d'estimations cognitives, de préférences, de proximités » (Cauvin, 1997 : 4).

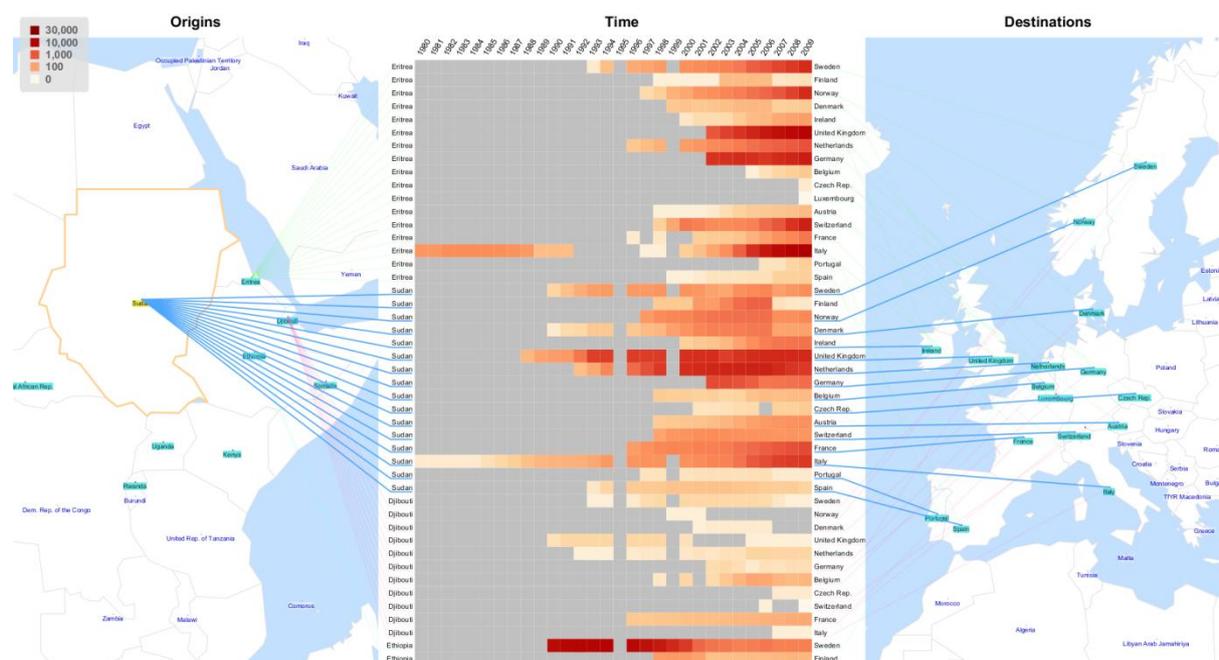
⁸¹ Nous mentionnons à titre d'exemple la représentation graphique habituelle du problème dit du marchand de glace, suivant la loi de H. Hotelling (1929), qui formalise l'espace de manière unidimensionnelle, sous la forme d'une ligne horizontale.

⁸² D'après l'encyclopédie collaborative en ligne Wikipédia© : « L'intérêt des coordonnées parallèles se trouve dans le fait que certaines propriétés géométriques de plusieurs dimensions se transforment facilement en 2D. Par exemple, une série de points sur une ligne dans n-plans se transforme en une série de lignes dans les coordonnées parallèles se faisant intersection à n-1 points. Lorsque n = 2, cela montre une dualité point ↔ ligne éclairant pourquoi les fondations mathématiques des coordonnées parallèles sont développées sur une projection plutôt que dans l'espace euclidien ».

des réseaux sociaux pour représenter les graphes bipartis⁸³, qui correspondent aux réseaux de relations où les sommets d'origine et de destination appartiennent à des catégories thématiques distinctes (les « *two modes networks* »).

Ce mode de représentation est théoriquement mis en œuvre dans le cadre d'un paradigme de la représentation orienté sur les lieux, mais nous pensons qu'il peut également convenir pour représenter des relations entre les lieux, comme l'illustre la Figure 2-31.

Figure 2-31. Une double carte de flux combinée à une matrice ordonnée



Source : Boyandin I. et al. (2011)

A noter que la version des coordonnées parallèles présentée sur cette figure présente l'intérêt de combiner une double carte de flux avec une matrice ordonnée temporelle.

Le formalisme de la coordonnée parallèle présente plusieurs avantages. Le premier est celui de dissocier les lieux d'origine et de destination en les plaçant sur deux axes séparés, la valeur du flux est portée sur un troisième axe. On notera toutefois que l'espace ayant recours, à lui seul, à deux axes : l'un pour les *x* et l'autre pour les *y*, deux variantes de ce modèle appliqué à la cartographie sont envisageables. Le second avantage des coordonnées parallèles est lié aux possibilités de visualisation de plusieurs caractéristiques des flux lorsque le nombre d'axes est multiplié. En effet, les axes peuvent faire l'objet d'une sélection ou d'un ré ordonnancement, sur le principe de la matrice ordonnée, afin de faire apparaître une structure particulière qui sera ensuite portée sur l'axe central, celui qui permet de faire le lien entre les deux axes des lieux d'origine et de destination. C'est la solution qui a été choisie par I. Boyandin et al. (2011) pour cartographier les flux de réfugiés présentés sur la Figure 2-31.

- Le « *kriskogramme* »

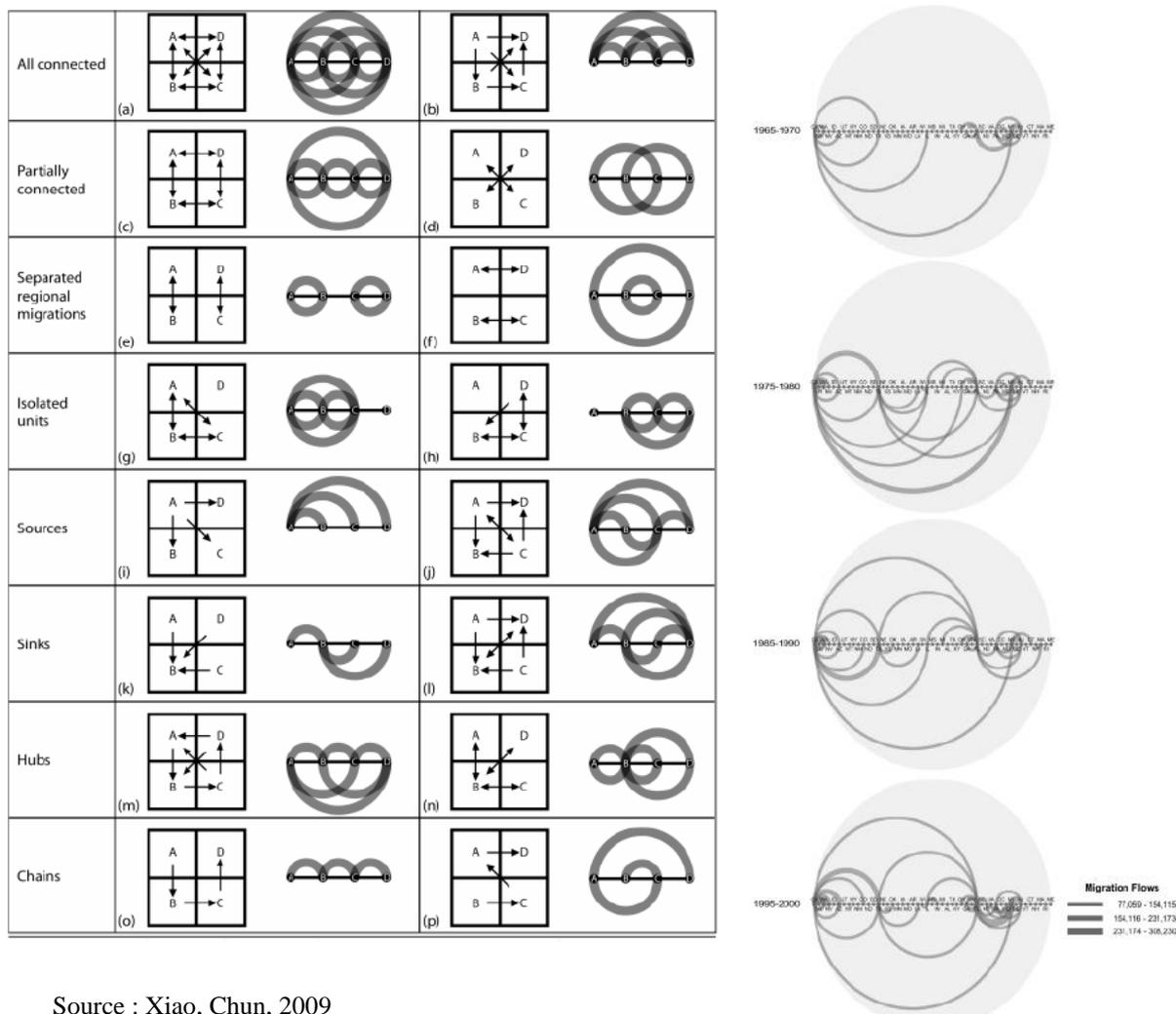


Cette variante de carte de flux a été développée récemment par (Xiao, Chun, 2009), sur le modèle des diagrammes en arcs⁸⁴. Le « *kriskogramme* » consiste à projeter les unités spatiales

⁸³ Des variantes de ces « *two-modes* » sociogrammes sont disponibles dans la littérature, c'est le cas de la représentation de Galois (Freeman, White, 1993)

sur un segment de droite horizontal, la ligne d'emplacement, puis à relier les couples de lieux par des demi-arcs de cercles, comme illustré par la partie gauche de la Figure 2-32.

Figure 2-32. La représentation de flux sous la forme de kriskogrammes (Xiao, et al., 2009)



Source : Xiao, Chun, 2009

Sur un kriskogramme, la ligne d'emplacement (le trait horizontal) représente l'espace sur une dimension : les lieux sont ainsi placés de manière à respecter la distance qui les sépare. Les flux sont figurés par des demi-cercles dont le diamètre correspond à la distance entre les lieux ; leur taille est proportionnelle à leur valeur de flux (voir Figure 2-32 à droite). L'orientation des flux est appréhendée par leur cardinalité, ce qui n'est pas courant et rend la lecture peu intuitive : les flux d'Est en Ouest seront, par exemple, placés au-dessus de la ligne d'emplacement et les flux d'ouest vers l'est en dessous. Le sens des flux n'est donc que partiellement pris en compte : on peut se demander ce qu'il en est des flux orientés du Nord vers le Sud.

La position des arcs de part et d'autre de la ligne d'emplacement est ainsi susceptible de créer un biais visuel en fonction de l'orientation privilégiée. Deux demi-arcs de cercles ne sont pas supposés se chevaucher sauf au droit des lieux dont ils sont l'origine ou la destination, ce qui réduit de fait *l'effet spaghetti*.

Pour alléger la figure, il convient de noter que les flux ont été présélectionnés en amont de la représentation. Les auteurs indiquent avoir eu recours à des méthodes fondées sur le tri des flux, selon certaines propriétés générales telles que l'orientation (est-ouest ou nord-sud, par exemple), ce qui permet de ne représenter que des demi-arcs de cercles, ou bien une sélection d'entre eux. Les auteurs ont aussi appliqués une procédure de réduction des données par le biais d'une classification, afin de limiter le nombre de figurés.

La méthode du Kriskogramme a été appliquée à une matrice de flux inter-états (48*48) des Etats-Unis disponible à quatre périodes (1965-1970, 1975-1980, 1985-1990, et 1995-2000). Les résultats obtenus (voir Figure 2-32 à droite) ont ensuite été comparés avec ceux d'autres auteurs, pour mettre en évidence par exemple la structure est-ouest des flux internes proposée dans un article de D. Plane (1999).

2.3.4.5 La transformation des liens

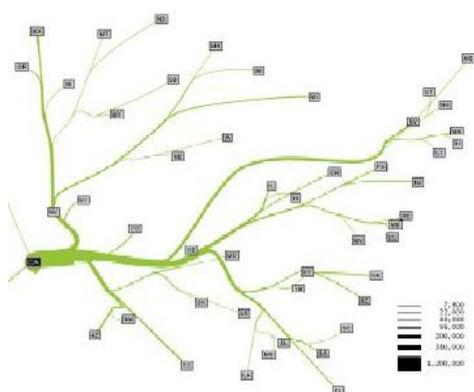


Deux variantes intéressantes de la carte de flux, fondées sur le principe de la transformation des liens, ont été proposées ces dernières années. Elles consistent à fusionner graphiquement les liens pour représenter, soit une sélection de flux sous la forme d'un arbre hiérarchique (Phan, 2005), soit l'ensemble des liens (*l'edge bundling*), à la manière de couloirs d'écoulement préférentiel (Holten, 2006).

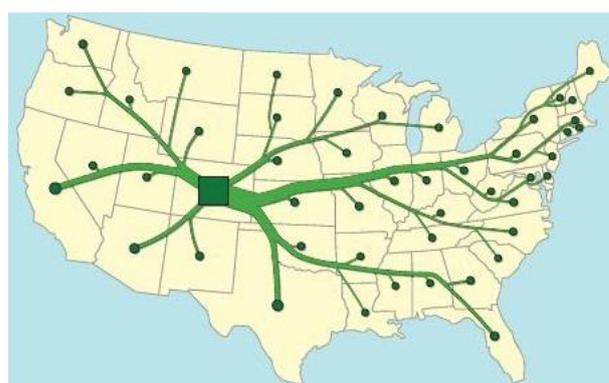
La Figure 2-33 illustre deux versions du mode de représentation des flux, sous la forme d'arbre hiérarchique. Leur application aux flux de l'ex-Tchécoslovaquie est illustrée par la Figure 2-34.

L'idée sous-jacente à ces deux méthodes consiste à agréger visuellement les lignes, en respectant au mieux la topologie des lieux, de manière à ce que les flux donnent l'apparence de converger dans certaines directions préférentielles. D'après K. Verbeek et al. (2011), la méthode permet de réorienter les flux en utilisant des routes intuitives. Elle consiste, en pratique, à faire comme si les données de flux – dont on rappelle que le cheminement exact entre deux lieux n'est pas connu – empruntaient des routes particulières formées par le maillage des lieux, à faire comme si elles étaient affectées sur un réseau planaire hiérarchique.

Figure 2-33. Illustrations d'une représentation des flux sous forme d'arbre hiérarchique



Source : Phan, D. et al. (2005)

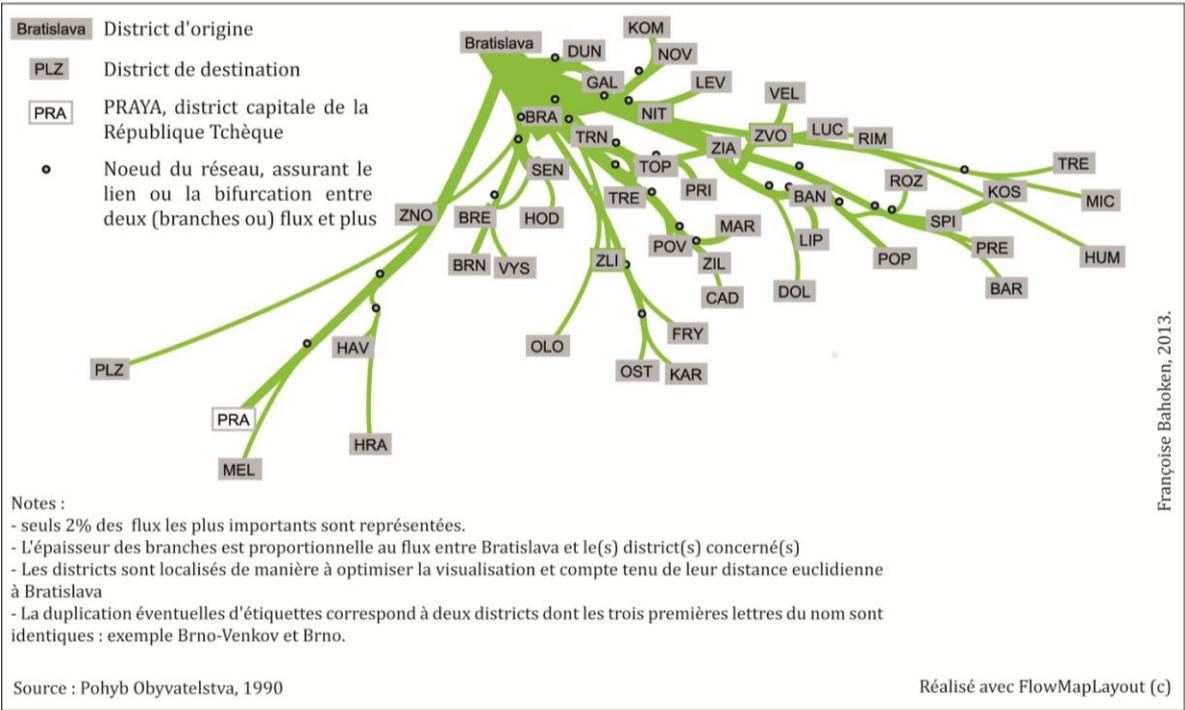


Source : Verbeek, K., et al (2011)

La version proposée D. Phan et al. (2005), applicable grâce à l'outil *Flowmap Layout* disponible en libre accès, ne semble applicable, à notre connaissance, que sur une sélection de lieux ; son efficacité semblerait surtout acquise sur des matrices creuses. C'est pourquoi la Figure 2-34 qui applique ce

procédé aux flux de l'ex-Tchécoslovaquie ne présente que ceux émis par le district de Bratislava, décrivant ainsi un graphe hiérarchique plutôt qu'une carte.

Figure 2-34. Un graphe des flux au départ de Bratislava (ex- République Tchèque), en 1990

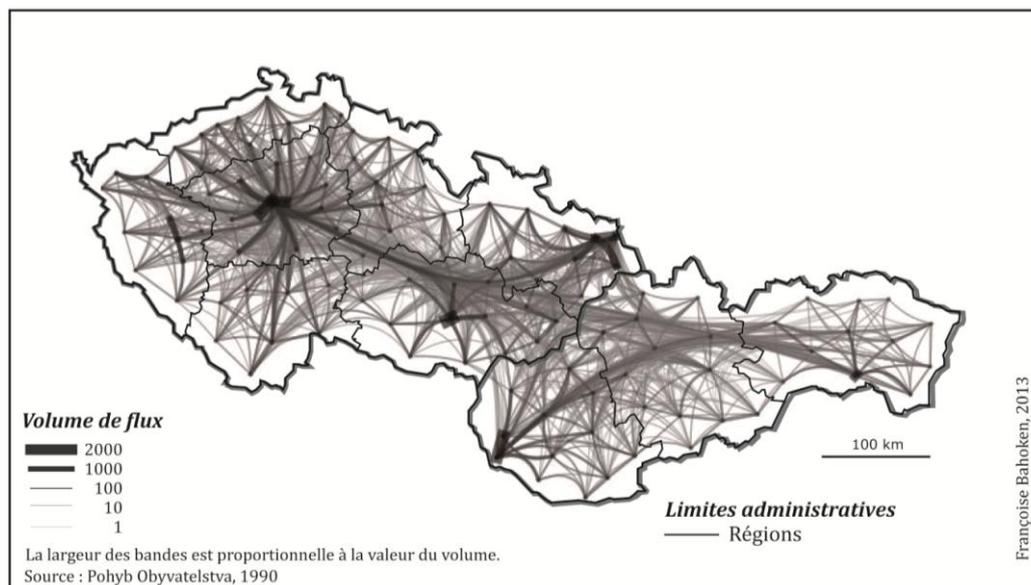


Sur la Figure 2-34, les nœuds sont placés automatiquement de manière à optimiser leur perception et l'importance de leur relation avec le point d'origine. Deux caractéristiques sont notables sur le plan graphique : à mesure que l'on s'éloigne du point d'entrée, le lien sortant du point d'origine se scinde en plusieurs bras, qui correspondent aux différentes destinations de ses flux, il s'affine aussi en fonction de l'éloignement des lieux de destination. Cette manière de représenter les flux fait écho à une pratique ancienne, celle des cartes de C.-J. Minard (1844), qui consiste à fusionner visuellement les flux au droit des lieux d'origine et de destination. Nous y reviendrons dans la partie consacrée à l'émergence de ce procédé graphique, dans la première section du chapitre 4.

La proposition de D. Holten (2006, 2009) va plus loin que la précédente, en permettant la cartographie de l'ensemble des liens qui illustrent les flux. Très connue aujourd'hui dans le milieu de l'informatique graphique, sous l'expression d'« *edge bundling* », la méthode consiste à fusionner visuellement le tracé des liens de manière à ce que *toute l'information disponible dans la matrice* soit représentée, et qu'elle mette en évidence des motifs des flux. Il est important de noter qu'il s'agit d'une avancée majeure dans le champ de la géovisualisation des matrices puisqu'il est dès lors possible de représenter tous les flux sur une carte, sans les sélectionner *a priori*.

L'application de cette méthode de fusion graphique des liens, pour cartographier le volume bilatéral des flux de l'ex-Tchécoslovaquie (1990), est présentée sur la Figure 2-35. L'image de ces flux qui en résulte met en évidence des routes préférentielles d'écoulement, qui correspondent à la fusion graphique des liens les plus faibles avec ceux qui présentent les valeurs les plus importantes, compte tenu de leur orientation. Il est important de noter que l'agrégation des liens est exclusivement graphique, elle intègre toutefois la valeur des flux ainsi qu'une contrainte de force-directe sur la position des lieux qui correspond au volume de flux mesuré aux lieux.

Figure 2-35. Cartographie du volume de flux de l'ex-Tchécoslovaquie (1990), selon le procédé de l'*edge bundling* hiérarchique



Ces méthodes qui sont développées dans le champ de l'informatique graphique présentent un intérêt certain pour la cartographie statistique de flux. Aussi allons-nous-y revenir en seconde partie (chapitre 5), certaines d'entre elles étant susceptibles de renouveler l'image classique des flux présentée sur la Figure I-1 (page 13).

Conclusion du chapitre

La représentation graphique des matrices de flux est l'une des difficultés majeures de la représentation graphique de données statistiques et cela, quelle que soit la forme de la figure et l'objectif principal de sa mise en œuvre.

La matrice ordonnée permet de mettre de l'ordre dans les données de la matrice de flux, en explorant puis en classant ses valeurs de manière à obtenir une typologie révélatrice de la structure des flux. En prenant la forme matricielle, elle conduit à la transformation des valeurs des flux en une trame de colorée ou de niveaux de gris. Elle revêt cependant deux inconvénients principaux qui limitent sa puissance exploratoire : le premier est qu'elle n'autorise pas une lecture aisée des relations locales et le second, quelle ne spatialise pas les flux. De ce point de vue, le graphe présente alors l'avantage de représenter la structure des données dans un plan, même s'il est fortement soumis à *l'effet spaghetti* : la matrice ordonnée n'étant, de fait, pas concernée par ces contraintes de complexité graphiques inhérentes aux autres modes de représentation des flux dans un plan.

La réalisation d'un graphe conduit à une double transformation de la matrice de départ : la première correspond à une modification de la figure et la seconde, à un changement du formalisme des valeurs. La modification de la figure correspond au passage de la forme matricielle à celui du *lien-nœud* tandis que la modification du formalisme des valeurs correspond à leur transformation en objets graphiques, essentiellement en points et en lignes. Cette double transformation conduit à la représentation de la structure des flux dans un plan, ce qui permet la visualisation des liens existants entre des lieux, éventuellement de leur qualité et/ou de la quantité de flux correspondante. Le principal inconvénient du graphe est qu'il soit généralement a-spatial. De ce point de vue, nous suggérons d'appréhender le graphe plutôt comme une figure complémentaire de la carte de flux, dans le cadre de la visualisation rapide de les résultats issus de l'exploitation des données, ou dans un objectif exploratoire de recherche de structure spatiale, plutôt que comme le produit final d'une analyse dite géographique.

La carte de flux correspond, en effet, formellement à la spatialisation du graphe, c'est-à-dire à la projection de la structure des flux dans un espace géographique. La nécessité d'une métrique constituant une contrainte forte supplémentaire qui tient à la position des lieux – fixe dans un système de projection – et à la notion de séparation qui lui est associée. Lorsqu'elle est réalisée à partir d'une matrice de flux dense, la carte – tout comme le graphe de flux – souffre d'une complexité graphique qui tient à la manifestation conjointe d'un effet de surcharge graphique et de couverture spatiale, tous deux conduisant à *l'effet spaghetti*. La solution classique apportée est mise en œuvre lors du traitement des données, elle intervient donc, en pratique, en amont de l'étape de représentation proprement dite. Nous verrons dans le chapitre 5 qu'il est possible de réduire la complexité graphique d'une carte de flux, en mobilisant les seuls aspects graphiques et sémiologiques.

Cette contrainte graphique se voit renforcée par le caractère multidimensionnel des données qui portent sur des flux. Leur représentation graphique constitue d'ailleurs, toujours, l'une des difficultés majeures du champ de la représentation de données statistiques. G. Palsky a noté qu'un des éléments d'explication courant aux XVIII et XIX^{ème} siècles était le fait que « *La carte traduit l'espace et le diagramme le temps* » (Palsky, 2004a : 346), d'où la difficulté de représenter des mouvements. L'auteur explique que cette position n'est plus valable, notamment après 1800, puisque plusieurs formes de cartographie inscrivent le temps dans l'espace, parmi lesquelles les cartes de flux.

La principale caractéristique de la carte de flux est qu'« [...] *elle spatialise le temps, mais de façon trompeuse, puisqu'une section courte ne signifie rien quant à la durée ou la vitesse de déplacement* ». (Palsky, 2004a : 347). C'est notamment pour cela que les procédés d'anamorphose apparaissent les plus adaptés à la représentation des changements spatiaux et qu'ils ont été qualifiés par S. Rimbart (1990) de « *stato-dynamique* ». Ils présentent cependant l'inconvénient d'être à la fois difficiles à mettre en œuvre et à lire, pour un public non averti.

La difficulté de son énoncé, de l'identification de ses causes et par voie de conséquence, des solutions à apporter au problème de *l'effet spaghetti* conduit encore aujourd'hui, nombre d'étudiants, de praticiens et de chercheurs à éviter de réaliser une carte de flux ou, du moins, à ne la réaliser que du point de vue des seuls lieux, ce qui consiste à réduire le nombre de ses dimensions. P.-A. Rosental (1997) a noté, à juste titre, que ces seules difficultés justifient le recours quasi exclusif au tableau statistique et au diagramme plus qu'à la carte de flux dans les travaux qui portent sur l'analyse quantitative des mobilités spatiales.

Le besoin de cartographier des flux de toutes natures n'est pourtant pas récent et différents auteurs ont tenté de gérer ce problème de complexité graphique. D. Guo (2009) mentionne l'urgence de nouvelles approches de visualisation cartographique des matrices qui permettraient de mieux comprendre la structure spatialisée des flux. Sa proposition est toutefois focalisée sur les seuls aspects liés à l'exploration et à la visualisation graphique, elle ne porte pas spécifiquement sur les aspects théoriques cartographiques.

Pour tenter d'y remédier, il est utile de connaître, avant toute chose, l'éventail des solutions cartographiques qui sont aujourd'hui disponibles pour cartographier des matrices de flux. Dans les deux chapitres suivants, nous examinerons la manière dont on cartographie le cœur de la matrice, c'est-à-dire les liens, les méthodes relatives à la représentation des flux du point de vue des lieux étant supposées bien connues. Nous proposerons dans le chapitre 3 une typologie des principales méthodes de cartographie de flux qui s'appuie sur des principes de décomposition de la matrice de données initiales (Fij). Dans le chapitre 4, nous examinerons plutôt la manière dont une sélection d'auteurs a réalisé des cartes de flux à travers le temps : cette rétrospective des méthodes de cartographie de flux sera accompagnée d'une présentation des différents procédés de représentation cartographique des matrices de flux, elle aura également pour objectif de valider la typologie proposée dans le chapitre précédent.

3^{ème} CHAPITRE. La décomposition des matrices de flux (Fij) : le volume (Fij⁺) et le solde (Fij⁻) bilatéral

Les matrices de flux correspondent à « [...] ces tables [qui] peuvent être décomposées en deux parties, une partie symétrique et une partie antisymétrique. Pour les statisticiens de l'auditoire la variance totale peut également être divisée en ces deux parties ».

Waldo Rudolf Tobler (1979).

« Lorsque l'on veut tout représenter sur une même carte [de flux], des simplifications s'imposent. Elles peuvent être de diverses natures : 1° Représentation de la balance entre les deux sens, ce qui divise le réseau par deux et évite les allers et retours. Mais la vision de la masse migrante disparaît [...] ».

Jacques Bertin (1973).

Introduction du chapitre

L'analyse cartographique des matrices de flux intéresse les méthodes de traitement de données cartographiques et numériques. En pratique, elle est d'abord focalisée sur l'exploitation des valeurs de la matrice avant leur représentation cartographique, en supposant que les données cartographiques l'ont été précédemment. L'ensemble de la procédure formant le corpus des méthodes de cartographie de flux. Nous présentons dans ce chapitre les principales méthodes qui forment ce corpus, afin d'aboutir à une typologie de celles qui constituent, de notre point de vue, le fondement de l'approche usuelle. Les méthodes liées à la construction du fond de carte ne seront pas présentées, car elles appartiennent au domaine de la cartographie statistique générale : celle des flux ne constituant pas un cas particulier. Les méthodes dont il est question s'inscrivent dans le cadre de raisonnements qui conduisent à la construction des valeurs de la matrice, c'est-à-dire à la modification des données initiales, avant leur représentation cartographique.

Il est important de noter que, quelle que soit la méthode de traitement utilisée pour générer les valeurs de flux, l'étape de cartographie consiste toujours à représenter l'une des trois familles de matrices sur laquelle nous allons aboutir à la fin de ce chapitre. Etant donné que la formation de ces matrices s'inscrit dans plusieurs approches du traitement des données, qui peuvent être descriptives, explicatives ou prédictives, nous restreignons volontairement notre propos à l'approche descriptive. Cela signifie que les valeurs disponibles dans notre matrice, pour la cartographie, résultent d'un traitement statistique simple, à visée descriptive. Ce traitement conduit à l'élaboration d'une nouvelle matrice, issue des données observées (de leur transformation, de leur réduction, de leur modélisation,...), dont les valeurs feront l'objet d'une représentation directe sous la forme de carte de flux.

Le point de départ de l'étape de cartographie statistique correspond toujours à la préparation du tableau numérique, de la matrice. De ce point de vue, la restriction de cette section à la présentation des principales méthodes de cartographie descriptive des flux n'en est pas vraiment une : les méthodes que nous présentons sont également applicables aux matrices issues des autres familles de traitements de données. **D'un point de vue cartographique, le fait que la matrice dont on dispose pour l'analyse résulte de telle ou telle famille de traitement importe finalement peu puisque, sur le plan formel, elle est formée d'une série de valeurs généralement quantitatives qui ne diffèrent que par leur signification (chapitre 1), et par l'interprétation de ce qu'elles sont censées décrire (nous confirmerons cela dans le chapitre 6).** Prenons un exemple : une matrice du volume de flux décrit la quantité qui s'est déplacé entre deux lieux, quand une matrice des résidus positifs décrit les relations préférentielles exprimées entre deux lieux, c'est-à-dire la seule sélection des quantités transférées entre deux lieux qui répond aux hypothèses de départ. Ainsi, ce qui compte de notre point de vue, au moment de l'étape de représentation cartographique de flux, c'est davantage la connaissance du contenu de la matrice, c'est-à-dire de *l'information disponible* pour réaliser la carte, de son existence pratique, que des traitements dont elle a fait l'objet.

J. Bertin suggère de cartographier d'une part, les liaisons orientées et d'autre part les flux (Bertin, 1973 : 344), soit de réaliser deux cartes distinctes⁸⁵. Procéder ainsi conduit à une dégradation de la matrice de données initiales (Fij) en une matrice de liens (Lij). La présentation de cette procédure dans la première section nous conduit à la notion de *signification* de cette *information disponible* pour la cartographie. Elle est appréhendée dans un double mouvement qui intéresse les niveaux théoriques et sémiologiques des flux, autrement dit la manière dont le flux, qui correspond nécessairement à un type de matrice, est exprimé ainsi que celle dont il est représenté. Dans le premier cas, il s'agit d'identifier l'information de flux (Fij) que l'on souhaite cartographier. *S'agit-il de fréquences, de probabilités d'émission ou bien de quantités ? D'un volume de flux ?*, par exemple. Dans le second cas, c'est plutôt la manière dont ce flux identifié est représenté pour pouvoir être perçu, c'est sa sémiologie graphique qui importe : *l'information analysée est-elle représentée par un système de doubles flèches ou non ? Doit-elle traduire une relation d'ordre ou non ?*

La deuxième section de ce chapitre présente deux des principales familles de traitement des données de flux initiales (Fij), celles qui répondent à la question de l'identification de l'information. Nous présenterons d'abord les analyses fondées sur les fréquences des valeurs de flux, puis l'approche matricielle proposée par les travaux de W. Tobler, qui conduit à la décomposition de la matrice de flux initiale en deux composantes complémentaires.

Dans la troisième section, nous repartons de cette proposition de W. Tobler de décomposition de la matrice (laquelle correspond d'ailleurs à celle de J. Bertin de simplifier l'information), en ne représentant que la balance migratoire. Il est intéressant de noter la proximité de ces deux approches, apportant ainsi une réponse à la seconde question : *quelle cartographie des flux est envisageable, en fonction du type de matrice ?* Nous proposons pour cela une typologie des méthodes classiques de cartographie de flux en trois grandes familles. La présentation est focalisée sur les aspects sémiologiques du cœur de la matrice : (Fij), les aspects concernant la sémiologie des lieux (i) ou (j) étant, comme précédemment, supposés bien connus. Ces méthodes intéressent le mode de représentation discret des flux qui mobilise, d'après W. Tobler (1982), des flèches ou des traits pour cartographier les données de flux, qu'elles soient initiales (élémentaires), décomposées ou bien transformées, en probabilités de migration par exemple. De fait, les modes de représentation des flux sous la forme de champs de vecteurs ne seront pas examinés, l'auteur ayant bien décrit leur mise en œuvre dans sa présentation des possibilités de la cartographie de flux (Tobler, 1982).

⁸⁵ Les cartes qui illustrent ce propos sont présentées sur la Figure 4-17, page 165.

Les trois familles de méthodes de cartographie de flux que nous présentons constituent, de notre point de vue, le socle de la cartographie classique de flux, son fondement : elles sont illustrées à nouveau par la matrice de flux de l'ex-Tchécoslovaquie (1990). La présentation de ces méthodes revêt un caractère essentiel dans la mesure où nous nous appuyerons sur celles-ci, dans la seconde partie de la thèse, pour proposer des solutions à une sélection de problèmes de la cartographie de flux.

3.1 *L'information disponible dans la matrice : ce que l'on peut cartographier (Lij)*

L'*information disponible* dans une matrice de flux quelconque fait référence à son existence pratique pour l'étape de cartographie. Elle concerne les valeurs de flux (F_{ij}) qui sont décrites par des liens (L_{ij}), en l'occurrence valués, ainsi que par les marges de la matrice (O_i, D_j) que nous avons présentées dans le chapitre 1 ; ces deux informations étant analysées conjointement. Cette posture qui consiste à ne pas dissocier la cartographie des flux de celle (de l'effet des flux au niveau) des lieux fait référence au cadre théorique proposé par C. Grasland (2009), à savoir l'analyse conjointe des interactions et des positions.

3.1.1 *L'information disponible dans la matrice*

Nous appréhendons la matrice dans le cadre de sa visualisation⁸⁶, au sens d'exploration, sous la forme d'un graphe spatialisé. En pratique, certaines cases (i, j) de la matrice sont renseignées par une valeur ($F_{ij} \neq 0$), d'autres ne le sont pas : leur valeur est *apparemment* égale à 0, mais ce n'est pas toujours le cas.

En fait, l'information disponible dans une matrice correspond aux cinq éventualités suivantes :

- cas n°1 des valeurs réellement non nulles : $F_{ij} \neq 0$;
- cas n°2 des valeurs réellement nulles : $F_{ij} = 0$;
- cas n°3 des valeurs non disponibles ($F_{ij} = \phi$) que nous fixons à 0, afin de ne pas disposer de cases vides : $F_{ij} = \phi = 0$;
- cas n°4 des valeurs inconnues ou aberrantes : $F_{ij} = ?$
- cas n°5 des valeurs impossibles.

Les cas n° 1 et 2 ne sont pas ambigus : il est, en effet, toujours possible d'identifier la Donnée dont il est question et son type, selon qu'elle est binaire, ordinale, quantitative discrète ou continue. Dans le cas n°2, une case de la matrice assortie d'une valeur réellement nulle ($F_{ij}=0$) signifie qu'il n'existe pas, dans la réalité, de relation entre (i) et (j). (F_{ij}) est dans ce cas une valeur impossible, à la différence de l'arrondi qui existe. La valeur 92, par exemple, lorsqu'elle exprimée en milliers correspond à 0,092, qui peut aussi être notée : 0, lorsqu'elle est arrondie. Plus généralement, une matrice telle que celle des flux aériens internationaux de passagers de l'OACI pour une année donnée dispose, par essence, de valeurs réellement nulles puisque tous les aéroports du monde ne disposent pas d'un vol à destination de tous les aéroports du monde entier. A l'inverse, les valeurs des cas n°3 et n°4 sont considérées, à défaut d'informations complémentaires, comme nulles : il s'agit de valeurs qui ont été éliminées, – volontairement ou non – par le producteur de la donnée ou par l'analyste.

Dans le cas n°3 ($F_{ij} = \phi$), la valeur est non disponible et correspond à l'ensemble vide, car elle est absente. Nous faisons l'hypothèse qu'elle a été éliminée volontairement et avançons quatre raisons

⁸⁶ La visualisation présente un aspect temporaire de la représentation graphique réalisée soit à des fins d'exploration du contenu d'un tableau, soit à des fins de communication. Les *visuels* sont, en effet, des supports temporaires, papiers ou numériques permettant de diffuser rapidement et instantanément une information généralement simplifiée, de transmettre un message particulier. Les visuels ne sont pas, de notre point de vue, supposés être pérenne : du moins, pas de la même façon que l'est une carte ou un graphique. C'est d'ailleurs la raison pour laquelle la mise en page des *visuels* scientifiques n'est que peu voire pas élaborée : les éléments classiques d'habillage et de légende y sont, dans bien des cas, absents.

principales pour expliquer cet appauvrissement de la matrice (des raisons qui relèvent autant du producteur de données que de l'analyste).

La première raison est imputable au producteur de la donnée : elle est liée à la granularité des filtres spatiaux-temporels appliqués par le dispositif de collecte sur la mesure du flux (voir chapitre 1). La donnée est donc inexistante parce que sa résolution est plus fine que les découpages du temps et de l'espace utilisés pour la mesure.

La seconde raison, également imputable au seul producteur de la donnée, tient à l'application de restrictions d'accès à l'information, pour des raisons de confidentialité ou d'ordre technique. Un premier cas de restriction est lié au fait que la précision sémantique et/ou la résolution spatiale est trop importante pour être mise à disposition du grand public : c'est le fait, par exemple, des informations issues du RP qui portent sur les flux d'étrangers. Ce caractère est appréhendé par la nationalité des individus nés étrangers à l'étranger et qui résident en France en tant qu'étranger. Ces données ne sont accessibles qu'à l'échelle du découpage territorial en IRIS de l'INSEE ou par agrégat de 5 000 individus au moins, en raison de leur sensibilité. Le second cas de restriction d'accès est d'ordre technique, il correspond aux capacités des outils d'archivage ou d'analyse de l'information. Le fichier *détail Mobilités professionnelles des individus* (MOBPRO) réalisé par l'INSEE pour l'année 2010 en constitue un bon exemple. Cette base de données sur des navettes quotidiennes est disponible à deux niveaux de précision : à l'échelle individuelle et à celle de flux agrégés. Ces deux matrices, issues de la même phase de collecte, présentent une quantité d'information qui varie en fonction du format d'archivage des données. Lorsque le fichier est fourni au format *.txt, il décrit⁸⁷ de façon exhaustive les déplacements rendus anonymes des 8 455 907 actifs de 15 ans et plus occupant un emploi, qui résident et travaillent en France (départements et territoires d'outre-mer inclus)⁸⁸ ou sur une commune transfrontalière. A l'inverse, le fichier disponible au format *.xlsx, pour Excel 2000© et versions antérieures, ne renseigne que sur les 26 476 valeurs agrégées de flux supérieures ou égales à 100 individus. D'après l'INSEE, la raison de la restriction de ce second fichier est liée au format d'Excel© antérieur à 2000 qui présente la contrainte des 65 000 lignes et des 255 colonnes. La restriction est donc liée aux capacités réduites de certains processeurs.

La troisième raison d'élimination volontaire de l'information tient à la mise en place de procédures liées à la confidentialité des données, parmi lesquelles l'arrondi des valeurs. Ces procédures sont appliquées par le producteur de l'information, parfois par l'analyste. Il est important de noter que l'effet de l'arrondi varie selon qu'il est réalisé lors du traitement des données statistiques ou au moment de l'étape de cartographie. L'arrondi à la valeur inférieure, ou l'extraction de la partie entière, d'une quantité continue contribue, en effet, à la sous-estimer tandis que l'arrondi à la valeur supérieure la surestime. Cette question, qui n'est pas à proprement parler cartographique, mérite toutefois que l'on s'y intéresse car elle agit de deux manières sur la carte : d'une part, en conditionnant la quantité d'information qui sera représentée et d'autre part, en influant sur la distribution des valeurs, sur leur répartition dans différentes classes lors de l'application de procédures de discrétisation. Cet arrondi rend l'information disponible incertaine et caractéristique d'un problème statistique de précision, d'incertitude des données que nous n'allons pas développer. L'effet de l'arrondi se manifeste non seulement sur les valeurs numériques – ce qui est un problème statistique – mais également sur les données cartographiques, sur le degré de précision des coordonnées géographiques dont la finesse

⁸⁷ Les fichiers de détail de MOBPRO sont accessibles via l'onglet Bases de Données au format *.txt (100 Mo) ; le fichier national contient 8 455 907 d'observations (26 684 772 observations après pondération) décrites par 32 variables ; le même fichier disponible au format *.xls (6 Mo) est accessible par l'onglet thématique.

⁸⁸ 98% (8 350 996) de ces individus résident et travaillent en France, hors communes transfrontalières.

n'est réellement justifiée⁸⁹ que dans certaines applications particulières. En effet, cette précision est bien souvent inutile étant donné l'échelle de représentation des cartes statistiques.

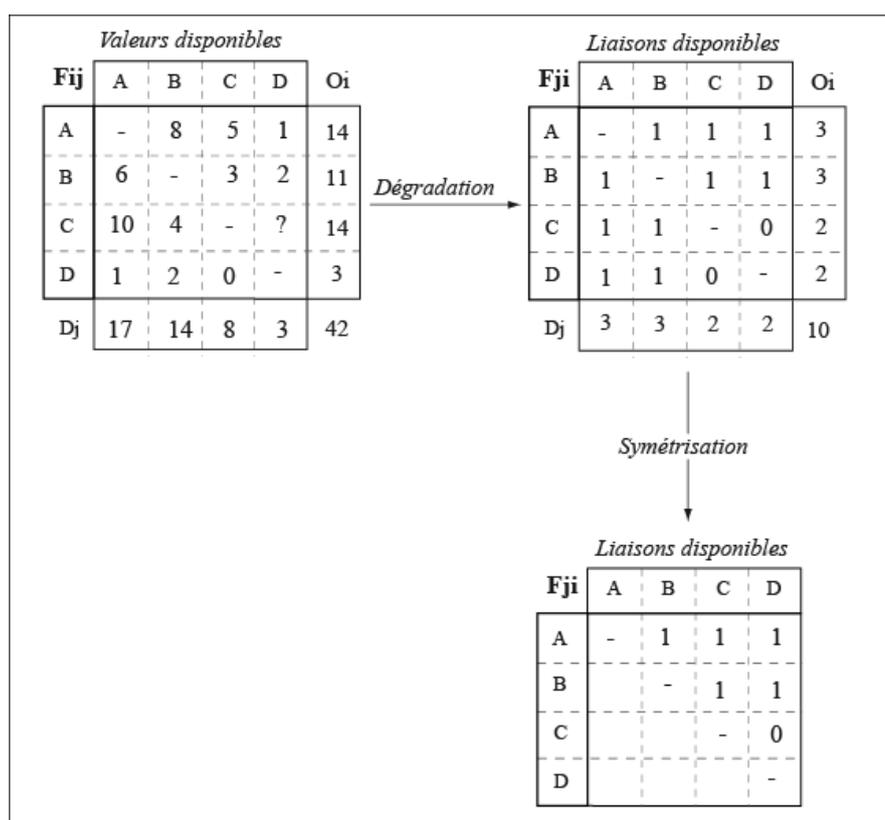
Le cas n°4 de valeurs aberrantes ou inconnues est très difficile à résoudre, car l'information est, justement, inconnue et l'on ne peut pas toujours savoir si une valeur apparemment aberrante ne correspond pas à un cas singulier, à l'instar des flux de migrants apparemment exceptionnels entre Toulon et Lorient qui sont simplement liés aux déplacements de marins entre arsenaux (Pumain, Saint-Julien, 1990 : 34).

Le cas n°5 correspond, quant à lui aux valeurs impossibles.

Ces cinq cas peuvent être regroupées en deux grandes familles: celle des valeurs impossibles et celle des valeurs possibles. Parmi ces dernières, on compte le cas des valeurs non disponibles, aberrantes ou non observées ($F_{ij}=?$), et le cas des valeurs disponibles : ($F_{ij} > 0$) ou ($F_{ij} = 0$).

Afin de ne pas alourdir l'analyse par ces considérations liées à la précision de l'information disponible dans la matrice, ou à leur caractère lacunaire, nous restreignons notre champ d'analyse aux cas n°1 et n°2. Pour visualiser l'information disponible, indépendamment de sa valeur, nous transformons légèrement notre matrice (F_{ij}) exemple, en introduisant une valeur nulle ($F_{ij}=0$) et une autre inconnue ($F_{ij}=?$), puis nous la dégradons⁹⁰ de sorte qu'elle devienne une *matrice d'adjacence* (voir Figure 3-1).

Figure 3-1. La transformation de la matrice de flux (F_{ij}) en matrice de liens (L_{ij})



⁸⁹ La précision hectométrique des coordonnées géographiques disponibles dans le système de projection RGF93, propres aux fichiers GEOFLA de limites administratives fournies par l'IGN, est de quatre chiffres après la virgule. Pour les applications courantes de cartographie statistique, cette précision nuit à la portabilité des données, en même temps qu'elle est inutile aux niveaux de résolution spatiale à laquelle ces coordonnées sont généralement utilisées.

⁹⁰ Cette dégradation correspond formellement à une binarisation : les cases renseignées de la matrice de flux prennent alors la valeur '1', sinon '0'. Cette méthode de transformation est généralement mise en œuvre sur des matrices valuées destinées à être exploitées à l'aide de méthodes issues de l'analyse des réseaux sociaux (*Social Network Analysis*) où l'existence d'une relation entre deux entités prime leur valeur. C'est parce que dans une matrice de flux, la valeur du flux prime sur son existence, que nous considérons que la binarisation est une dégradation de la matrice.

Cette matrice des liaisons (L_{ij}) ne décrit que des liens possibles (L_{ij}), qu'ils soient connus ou inconnus, nuls ou non. De rang N , cette matrice est désormais formée de $N(N-1)$ cellules assorties de valeurs binaires, qui prennent la modalité '1' s'il existe une valeur dans une cellule, c'est-à-dire une relation symétrique entre (i) et (j), sinon '0'. Nous avons en effet supposé que cette matrice est également complète s'il existe un lien entre (i) et (j), alors il existe le même lien un entre (j) et (i), c'est pourquoi nous l'avons symétrisée. Cette matrice n'est cependant pas orientée.

C'est donc l'ensemble des liaisons qui sont en pratique disponibles – qui ne correspondent pas nécessairement aux liaisons théoriques – qu'il est possible de représenter sur une carte. La visualisation d'une telle matrice de liens (L_{ij}), même dense, sous la forme d'un graphe ne pose pas de problèmes particuliers, comme en témoigne la quantité de travaux réalisés notamment dans le champ de la visualisation des réseaux sociaux. Ce qui pose en revanche problème, c'est la visualisation de cette même matrice sous la forme d'une carte de flux, en raison de la composante spatiale de localisation des lieux (voir section 2.3.3.1, page 86, dédiée à la présentation de *l'effet spaghetti*).

La Figure 3-2 illustre la représentation sur une carte de l'information disponible dans la matrice prise pour exemple. Celle-ci n'étant pas représentative du problème de complexité graphique, nous présentons également sur la Figure 3-3 la matrice des flux interdistricts tchécoslovaques.

Figure 3-2. Illustration de la cartographie de l'information disponible (L_{ij})

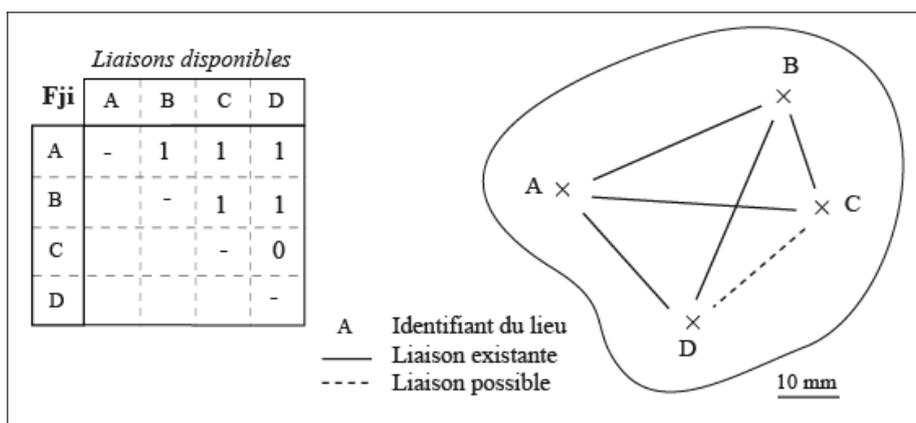
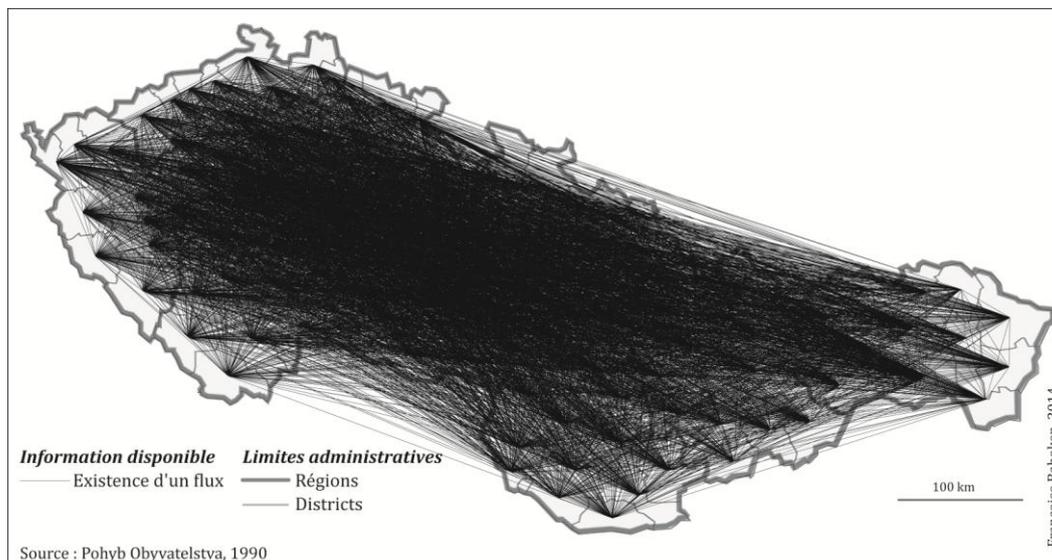


Figure 3-3. Cartographie de l'information disponible (L_{ij}) dans la matrice de flux de l'ex-Tchécoslovaquie (1990)



Cette forme de représentation des seules relations existantes (Lij) caractérise une approche que nous qualifions de *logique de lien*. Elle conduit à une allure particulière de carte de flux qui ne conduit pas à la représentation des valeurs des flux. La représentation de cette matrice (voir Figure 3-3) s'apparente, en réalité, à la spatialisation d'un graphe : la matrice étant très dense, il ne sert à rien de la cartographier car elle est soumise à **un effet spaghetti** maximal. Cet effet étant essentiellement lié aux problèmes de *couverture spatiale* (les liens n'étant pas valués).

Un élément supplémentaire d'explication, lié au découpage territorial spécifique à la matrice des flux de l'ex-Tchécoslovaquie, peut être avancé. Il tient à la configuration des zones dont la superficie est d'autant plus petite que l'attractivité des lieux est importante : les districts des deux capitales ou des trois républiques historiques polarisent, en effet, fortement les flux, alors que leur superficie est peu importante, relativement à celle des districts qui les avoisinent. Ces districts sont aussi ceux dont le nombre de degrés est important, c'est-à-dire qu'ils sont en relation effective (possible, disponible, non nulle) avec de nombreux autres districts, ce qui entraîne une superposition des liens importante.

Pour autoriser une représentation de la matrice lisible et *significative*⁹¹ sous la forme d'une carte de flux, l'approche générale consiste à appliquer un critère de sélection des valeurs à représenter et/ou une réduction des données en amont de leur représentation. L'application de critères de sélection faisant l'objet des chapitres 8 et 9, nous présentons dans la section suivante les principales méthodes de cartographie statistique descriptive de flux, sans nous préoccuper de la manière dont les valeurs de flux sont sélectionnées.

L'approche statistique descriptive consiste à transférer les méthodes d'analyse des tableaux de contingence aux matrices de flux (Beguin, Pumain, 2000). Elle a aussi conduit à l'élaboration de toutes sortes d'indices qu'il n'est pas envisageable de présenter de manière exhaustive. Nous restreignons notre propos à la présentation de ceux qui nous apparaissent fondamentaux⁹². Ces méthodes répondent, en effet, de manière simple à l'objectif de réduction de la complexité analytique de la matrice de flux, la réduction de la complexité graphique passant par la sélection ou l'agrégation des données. Ces méthodes constituent aussi, de notre point de vue le socle d'une analyse quantitative à visée descriptive des mobilités spatiales. Etant donné que ces méthodes s'appuient sur l'analogie de la matrice avec le tableau de contingence, nous rappelons sa construction dans la sous-section suivante.

3.1.2 Des liens et des lieux : l'analogie de la matrice de flux avec le tableau de contingence

La matrice de flux est assimilable à un type de tableau statistique d'effectifs à double entrée : le tableau de contingence. Chaque case y représente la valeur de la relation entre deux caractères quantitatifs, généralement notés (X, Y). Cette relation est étendue à l'analyse des matrices de flux où le caractère X correspond au lieu d'origine (i) et le caractère Y au lieu de destination (j), tel que présenté sur la Figure 3-4.

⁹¹ Nous reviendrons sur la question de la *significativité* des valeurs de flux à cartographier dans le chapitre 8.

⁹² Les indicateurs courant de l'analyse des matrices de flux interprétées comme des réseaux de relations non spatialisées peuvent également être calculés, dans le cadre d'une analyse de type Analyse des réseaux sociaux (*Social Network Analysis*) où le lieu est assimilé à un acteur. C'est notamment le cas des indicateurs fondés sur le nombre de degrés (entrants, sortants, moyen, etc.) et des coefficients de proximité (*closeness*) ou encore d'intermédiarité (*betweenness*).

Figure 3-4. Du tableau de contingence à la matrice de flux

Tableau de contingence					Matrice de flux					Valeurs disponibles					
					Destinations										
					<i>i</i>	<i>j</i>	<i>j</i> ₁	<i>j</i> ₂	<i>j</i> _p	<i>O_i</i>					
Modalités des origines	<i>i</i> ₁	<i>n</i> ₁₁	<i>n</i> _{1j}	<i>n</i> _{1p}	<i>n</i> _{1.}										
	<i>i</i> ₂	<i>n</i> ₂₁	<i>n</i> _{2j}	<i>n</i> _{2p}	<i>n</i> _{2.}										
	<i>i</i> _k	<i>n</i> _{k1}	<i>n</i> _{kj}	<i>n</i> _{kp}	<i>n</i> _{k.}										
	<i>D_j</i>	<i>n</i> _{.1}	<i>n</i> _{.j}	<i>n</i> _{.p}	<i>n</i> _{..}										
					<i>i</i>	<i>j</i>	<i>j</i> ₁	<i>j</i> _n	<i>O_i</i>						
Origines	<i>i</i> ₁	<i>F</i> ₁₁	<i>F</i> _{1j}	<i>F</i> _{1n}											
	<i>i</i>	<i>F</i> _{i1}	<i>F</i> _{ij}	<i>F</i> _{in}											
	<i>i</i> _n	<i>F</i> _{n1}	<i>F</i> _{nj}	<i>F</i> _{nn}											
	<i>D_j</i>	<i>F</i> _{.1}	<i>F</i> _{.j}	<i>F</i> _{.n}											
					<i>F_{ij}</i>	A	B	C	D	<i>O_i</i>					
A	-	8	5	1	14										
B	6	-	3	2	11										
C	10	4	-	1	15										
D	1	2	3	-	6										
<i>D_j</i>	17	14	11	4	46										

Source : D'après, Groupe Chadule (1987), page 77 et suivantes.

Un tableau de contingence exprime, en référence au terme anglais conditional, « [...] ce qui conditionnel, ce qui est produit sous certaines conditions » (Chadule, 1987 : 78). Ainsi, sur la partie gauche de la Figure 3-4, le caractère X correspond aux modalités des lieux d'origine (i) et le caractère Y à celles des lieux de destination (j). Les modalités de chacun de ces caractères (X) et (Y) sont alors respectivement la liste des lieux d'origine (i₁, ..., i_n) et celle des lieux de destination (j₁, ..., j_n) de la zone d'étude.

Dans un tableau de contingence (voir Figure 3-4, partie gauche), la mesure est portée au croisement d'une ligne et d'une colonne noté (n₁₁) et dans une matrice de flux (F_{ij}), comme illustré par notre matrice fictive (voir Figure 3-4, partie droite). Dans les deux cas, le premier terme correspond à l'identifiant de l'origine et le second à celui de la destination. Chaque case (n_{ij}) indique l'effectif de la modalité (i) des origines et la modalité (j) des destinations ce qui, pour les flux correspond à l'effectif dont l'origine est (i) et la destination (j). Les effectifs en ligne (n_{k1}, ..., n_{kp}) forment la distribution conditionnelle des destinations si l'origine est (i). Inversement, les effectifs en colonnes (n_{1p}, ..., n_{pp}) forment la distribution conditionnelle des origines si la destination est (j).

La propriété principale d'un tableau de contingence est l'additivité en ligne et en colonnes, que nous avons sommairement présentées dans le chapitre 1. La somme des modalités d'une cellule (i_j) correspond alors aux sommes marginales en lignes (O_i) et en colonnes (D_j). Le terme (n_{i.}) correspond à la somme en ligne et (n_{.j}) à la somme en colonnes, avec :

$$n_{i.} = n_{i.} + \dots + n_{ij} + \dots + n_{ip} = \sum_{j=1}^n n_{ij} = O_i \tag{3.1}$$

$$n_{.j} = n_{1j} + \dots + n_{ij} + \dots + n_{kj} = \sum_{i=1}^n n_{ij} = D_j \tag{3.2}$$

Cette propriété d'additivité correspond, dans une matrice de flux (voir Figure 3-4 à droite), aux marges de la matrice formées par la somme des valeurs de flux ayant pour origine le lieu (i) et pour destination le lieu (j), c'est-à-dire au *volume des flux aux lieux* (V_i) déjà défini dans le chapitre 1. La correspondance entre les notations est alors la suivante :

$$F_{i.} = \sum_j F_{ij} \text{ pour la valeur totale des flux émis par } i \text{ (somme marginale en ligne)} = O_i$$

$$F_{.j} = \sum_i F_{ij} \text{ pour la valeur totale des flux reçus par } j \text{ (somme marginale en colonne)} = D_j$$

$$F_{..} = \sum_i \sum_j F_{ij} \text{ pour la valeur totale des flux de la matrice (sommes en ligne et en colonne)}$$

C'est cette propriété d'additivité interne propre aux tableaux de contingence qui, transférée au cas des matrices de flux, a permis de compléter la matrice (F_{ij}) initiale et de transférer les méthodes du calcul statistique à l'analyse des mobilités spatialisées.

L'approche statistique descriptive est, en effet, très utile dans l'analyse cartographique des flux, car il est très souvent préférable de représenter un indicateur *significatif* des valeurs de flux plutôt que les valeurs observées. Le champ de l'analyse statistique des données de flux dans un objectif de cartographie étant potentiellement important, nous avons fait le choix de présenter deux principales familles de méthodes, qui sont aussi les plus couramment mises en œuvre : elles font l'objet de la section suivante.

La première famille de méthodes repose sur l'analyse des probabilités de fréquences des tableaux de contingence appliquée au cas des matrices de flux, elle est présentée dans la première sous-section. La seconde famille s'appuie sur les principes de décomposition de la matrice initiale en composantes élémentaires, plus faciles à analyser. Suggérée par W. Tobler, cette seconde approche fait l'objet de la seconde sous-section. Afin de distinguer nettement ces deux grandes approches, dont les différences peuvent apparaître à première vue ténues, nous avons pris le parti de les qualifier respectivement d'approches statistiques et matricielles.

3.2 Les approches statistiques et matricielles du traitement des données de flux (F_{ij})

Si la cartographie de flux intéresse plusieurs approches qui peuvent être redondantes ou contradictoires (les auteurs n'étant pas toujours d'accord entre eux), nous avons fait le choix de nous intéresser aux deux principales familles du traitement statistique de données de flux, en préalable de leur cartographie.

La première approche est principalement mise en œuvre en démographie, dans le cadre de l'analyse quantitative descriptive des migrations humaines qui examine, notamment, les probabilités d'émission et/ou de réception des flux répartis dans des zones géographiques. Largement développée par D. Courgeau dès les années 1970, elle a ensuite fait l'objet d'une appropriation par plusieurs géographes de la population, parmi lesquels W. Tobler, qui s'y intéresse à la fin des années 1970. L'auteur propose en particulier la réalisation d'une cartographie des champs de probabilités de migration, d'émission ou de réception des flux, en trois dimensions (Tobler, 1981b). L'approche probabiliste est très populaire, elle a donné lieu à la production d'une multitude d'indicateurs fondés sur l'analyse des fréquences des valeurs que nous ne pouvons présenter de manière exhaustive (Ravenstein, 1885 ; Dorling, 1991 ; Tobler, 1995 ; Holland, Plane, 2001).

La seconde approche s'inscrit exclusivement dans le cadre des travaux menés par W. Tobler, que nous regroupons sous le terme d'approche matricielle, afin de les distinguer de l'approche probabiliste que nous considérons plus classique. L'approche fondée sur les probabilités ayant été largement reprise alors que celle de W. Tobler nous apparaît plus confidentielle. Cette dénomination d'approche matricielle n'est pas nôtre : elle est liée au fait que l'auteur dit emprunter à la théorie matricielle mathématique (et non à la statistique) des notions et des concepts qu'il transfère ensuite à l'analyse géographique des flux. Le résultat conduit à la décomposition de l'information disponible dans la matrice de flux initiale en deux composantes complémentaires, qui correspondent à deux sous-matrices et en pratique, à deux nouvelles matrices de flux. Nous nous appuyons sur ces matrices de flux dans la troisième section, pour présenter ce que nous considérons être les trois grandes méthodes de cartographie de flux, une typologie appréhendée sur le plan sémiologique.

3.2.1 L'approche statistique : la cartographie des probabilités de fréquence des valeurs de flux

Le raisonnement fondé sur une approche statistique probabiliste a pour objectif de comparer entre elles soit des sous-populations (i,j) de tailles variables, c'est-à-dire différentes matrices portant sur la même

zone, soit les données d'une même matrice disponibles à plusieurs dates. Le premier cas, qui est celui qui nous intéresse ici, entraîne une transformation des valeurs de la matrice en fonction de la manière dont les probabilités sont interprétées. Comme l'indique D. Courgeau (2002, 2006), plusieurs interprétations des probabilités sont, en effet, envisageables mais toutes ne sont pas utiles à l'analyse des flux.

D. Courgeau définit la probabilité comme « [...] une fonction numérique non négative et additive, définie sur une tribu⁹³ (collection de sous-ensembles), dont la masse totale est l'unité » (Courgeau, 2006 : 2). Deux voies d'interprétation des probabilités font sens dans un objectif de cartographie de flux : la première repose sur l'analyse des fréquences simples et des fréquences conditionnelles ; la seconde repose sur la notion de densité conditionnelle. La probabilité de fréquence est interprétée, soit comme la limite de la fréquence d'apparition d'un événement lorsque ce dernier est présent plusieurs fois dans la série observée : c'est la fréquence simple ; soit comme la fréquence d'apparition d'un événement (A), sachant un événement (B), notée : $P(A|B)$ pour la probabilité conditionnelle.

Appliquées à l'analyse d'une matrice de flux à cartographier, ces différentes interprétations des probabilités conduisent à générer de nouvelles matrices qui décrivent, soit des probabilités de fréquence simple, soit des probabilités de fréquence conditionnelles. Dans le cas d'une transformation de la matrice de flux initiale en un tableau de fréquences simples, la valeur résulte du calcul de pourcentages (%) en lignes (ou en colonnes) rapporté au total général des flux (F..). La matrice transformée décrit alors des fréquences de valeurs de flux rapportées au total général, tel qu'illustré par le Tableau 3-1 qui présente une sélection d'indicateurs de fréquences simples.

Tableau 3-1. Indicateurs de fréquence simple des flux

Indicateur : part (en %) des flux (O _i , D _j)	Formulation
- pour les lieux	
pour origine i dans le total général des flux	$(F_{i./F..}) * 100$
pour destination j dans le total général des flux	$(F_{.j}/F..) * 100$
- pour les flux	
pour origine i et pour destination j dans le total des flux	$(F_{ij}/F..) * 100$
pour destination j et origine i dans le total des flux	$(F_{ji}/F..) * 100$

Dans le second cas, la matrice est celle de fréquences conditionnelles [(F_{ij}) | O_i, D_j]. Si les fréquences conditionnelles sont généralement exprimées en proportions du total en ligne (F_{i.}) et / ou en colonne (F_{.j}), de nombreux auteurs ont suggéré de pondérer ces fréquences en tenant compte des populations des lieux d'origine et de destination (Ravenstein, 1885 ; Dorling, 1991 ; Tobler, 1995 ; Holland, Plane, 2001), c'est-à-dire de critères externes et non endogènes liés à la taille des lieux ou à leur attractivité.

Ainsi, pour les flux de population, C. Holland et D. Plane (2001) proposent l'utilisation de plusieurs indicateurs parmi lesquels la « probabilité de transfert » (*transfert probability*) (Holland, Plane, 2001 : 91-92) qui correspond au rapport du nombre de migrants sur la population de la zone d'origine : (F_{ij} / P_i) . Les auteurs introduisent également la notion « d'impact migratoire » (*migration impact*) calculé pour chaque couple de lieux qui appartiennent à un sous-système des flux (Holland, Plane, 2001 : 52 ; Henrie, Plane, 2008 : 51-52), ce qui suppose de raisonner à l'échelle de la zone d'étude, ainsi qu'à celle de sous-parties⁹⁴ de la matrice. D. Dorling (1991) suggère de son côté l'utilisation

⁹³ La tribu A d'un ensemble de valeurs (Ω) est un sous-ensemble de parties vides ou composées d'une suite finie d'entités et sur laquelle des opérations combinatoires ensemblistes s'appliquent. Ces possibilités sont l'union (U) et l'intersection (∩) de sous-parties.

⁹⁴ L'impact migratoire indique « [...] le changement dans la population (positif ou négatif) qui résulte du turnover migratoire entre Los-Angeles et les autres CBSA [le Core Based Statistical Area, qui correspond au découpage

complémentaire aux (P_i , P_j) du moment de la journée (diurne ou nocturne). Quel que soit le mode de calcul, les valeurs obtenues expriment la probabilité conditionnelle des lieux à émettre (ou à recevoir) des flux, compte tenu de leur capacité d'émission (ou de réception).

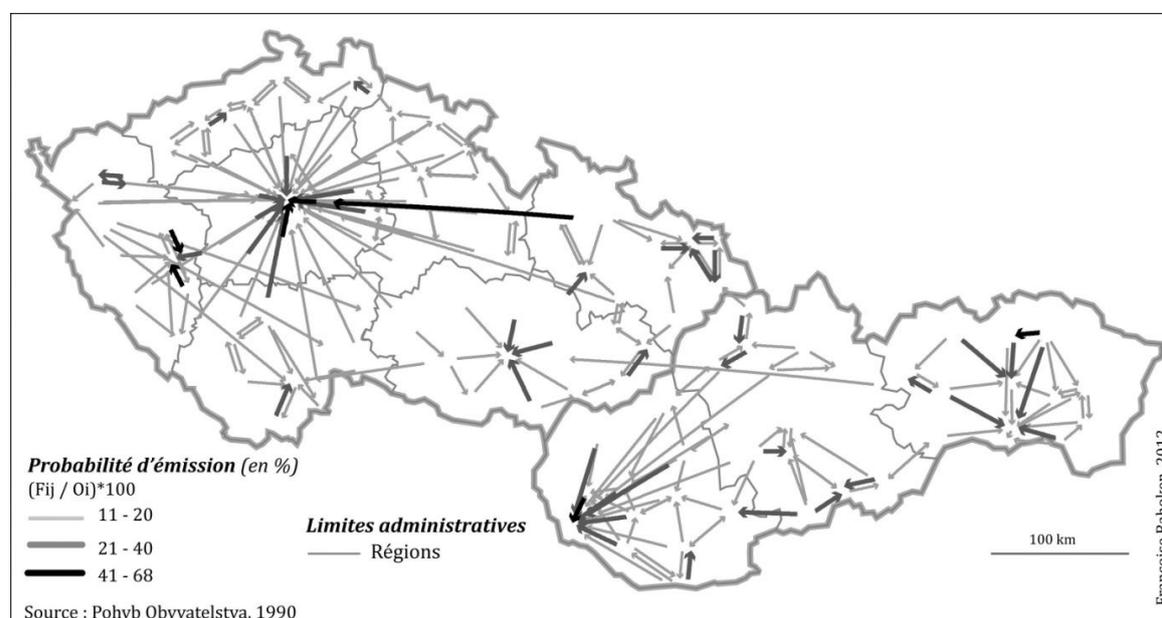
Le Tableau 3-2 présente une sélection d'indicateurs de fréquences conditionnelles qui peuvent être élaborés : il décrit des probabilités d'émission (ou de réception) des flux (F_{ij}) par des lieux d'origine (ou de destination) qui sont exprimées en utilisant les marges de la matrice (Figure 3-4, page 116).

Tableau 3-2. Indicateurs de fréquence conditionnelle

Indicateur : fréquence conditionnelle (en %) des flux ayant	Formulation
pour origine i et destination j / total des (i)	$(F_{ij}/F_{i.}) * 100$
pour origine i et destination j / total des (j)	$(F_{ij}/F_{.j}) * 100$
pour destination j et origine i / total des (j)	$(F_{ji}/F_{.j}) * 100$
pour destination j et origine i / total des (i)	$(F_{ji}/F_{i.}) * 100$
pour origine i et destination j / somme des marges	$F_{ij} / (F_{i.} + F_{.j}) * 100$
pour destination j et origine i / somme des marges	$F_{ji} / (F_{i.} + F_{.j}) * 100$
pour origine i et destination j / produit des marges	$F_{ij} / (F_{i.} * F_{.j}) * 100$
pour destination j et origine i / produit des marges	$F_{ji} / (F_{i.} * F_{.j}) * 100$

La Figure 3-5 présente un exemple de cartographie type d'indicateur de fréquence : elle illustre les probabilités d'émission $[(F_{ij}) / O_i] * 100$ des districts de l'ex-Tchécoslovaquie (1990). Il est possible de substituer les marges à des données exogènes qui sont fonction du type de flux : les populations des lieux d'origine et de destination pour les flux migratoires, par exemple.

Figure 3-5. Cartographie des probabilités d'émission (en %) de l'ex-Tchécoslovaquie (1990)



La transformation des valeurs initiales de la matrice en indicateurs de fréquences présente plusieurs avantages. Tout d'abord, elle intègre les valeurs des flux et leur effet marginal aux lieux : la

territorial américain utilisé pour la collecte de données statistiques] et en dehors des unités spatiales du CBSA avec lesquelles il existe un flux, pour 1000 personnes qui résident dans les autres (pas à Los-Angeles) CBSA ou en dehors des CBSA du comté. Il convient de noter que, à l'instar des taux de migration nette, il s'agit d'un indicateur de l'impact migratoire et non d'un taux auquel un événement démographique se produit, pour une population définie comme présentant un risque de vivre ces événements » (Holland, Plane, 2001 : 52).

représentation intègre alors l'ensemble de l'information disponible dans la matrice (liens, lieux) sans opérer de sélection des valeurs selon la focale de la matrice : *je représente les flux du point de vue des lieux* ou *je représente les flux entre les lieux*.

Le second avantage du recours aux fréquences est issu du premier : ces indicateurs permettent de mettre en évidence les parts d'échanges qui s'éloignent de la distribution marginale, autorisant ainsi l'analyse comparative de différentes séries.

Le troisième intérêt de l'utilisation des fréquences réside dans le calcul du seuil de densité (dk) de la matrice, qui représente le rapport entre la fréquence d'une classe et son amplitude.

Le quatrième avantage du recours à l'analyse des fréquences tient dans la possibilité d'extension du calcul pour l'estimation de sous représentations et de surreprésentations des effectifs dans la distribution totale. Pour cela, il est nécessaire de créer une troisième matrice qui représente l'écart entre les distributions des fréquences conditionnelles et les distributions marginales. Cette troisième matrice sera formée de valeurs respectivement positives (et négatives) qui représentent les surreprésentations (ou les sous représentations).

Se pose alors le problème de la diagonale vide qui fausse ces calculs de fréquences et les rends difficiles à exploiter.

Les quelques indicateurs de fréquences des valeurs de flux qui ont été présentés s'appuient sur la propriété d'additivité en ligne (et en colonne) de la matrice. Ils combinent généralement les valeurs de flux (F_{ij}) avec leurs marges (O_i , D_j) ou avec le total général des flux. Il est possible d'aller plus loin, en combinant les cases de la matrice qui renseignent sur les flux (F_{ij}) pour générer soit des indicateurs qui les décrivent – et non seulement les lieux – (Courgeau, 1980 : 84), soit de simplifier l'information de départ. Pour cela, il convient d'utiliser les méthodes issues du calcul matriciel qui ont été transférées à la cartographie de flux par W. Tobler.

3.2.2 L'approche matricielle

L'approche matricielle a été transférée à l'analyse des matrices de flux à la fin des années 1970 par W. Tobler, au moment où il travaillait sur la version additive du modèle gravitaire qu'il proposa par la suite.

L'auteur était, à l'époque, en train de résoudre « [...] *un problème pragmatique d'aménagement par une solution théorique très générale et très astucieuse* [...] » (Cauvin, Reymond, 1991 : 468) qui consiste à retrouver les échanges entre les lieux (F_{ij}), c'est-à-dire les valeurs des cases de la matrice, à partir d'une matrice des distances (D_{ij}) et des sommes marginales (O_i , D_j). W. Tobler en arrive alors à décomposer mathématiquement la matrice en sous-composantes élémentaires, en s'appuyant sur les principes du traitement matriciel.

L'idée de W. Tobler, d'utiliser le traitement matriciel, le conduit à démontrer que la meilleure façon de cartographier des flux est de représenter des champs de vecteurs⁹⁵, calculés sur ce qu'il nomme « *la composante négative des flux* », notée (F_{ij}^-). Cette réflexion sur les composantes négatives (F_{ij}^-) et par voie de conséquence positives (F_{ij}^+) des matrices de flux s'inscrit ultérieurement dans un courant de recherche plus important sur la synergie⁹⁶ qui analyse la dynamique des systèmes et leur stabilité (Sanders, 1992). W. Tobler poursuit cette idée générale en proposant de distinguer ce qui relève de la

⁹⁵ Ces champs de vecteurs devant le conduire à la notion d'attractivité.

⁹⁶ La synergie est un champ de recherche, développé à Stuttgart qui fournit un cadre théorique et méthodologique d'analyse de la dynamique de systèmes complexes, en particulier de modélisation de leur évolution. Les travaux qui y sont développés visent en particulier à simuler l'évolution de peuplements en agrégats, en accordant un rôle important à la stabilité du système analysé, sans pour autant en négliger la dynamique. On peut mentionner, à titre d'exemple, les travaux menés par L. Sanders (1992) sur l'évolution des systèmes de villes et ceux de Weidlich et Haag (1988) sur les migrations interrégionales.

dynamique du système de flux (F_{ij}^*), c'est-à-dire du changement dans l'espace, de ce qui concerne l'ossature du système (F_{ij}^*), sa stabilité. Pour cela, il s'appuie sur l'idée de décomposer les tableaux d'échanges en utilisant les propriétés de symétrie de la matrice.

3.2.2.1 Le caractère symétrique d'une matrice

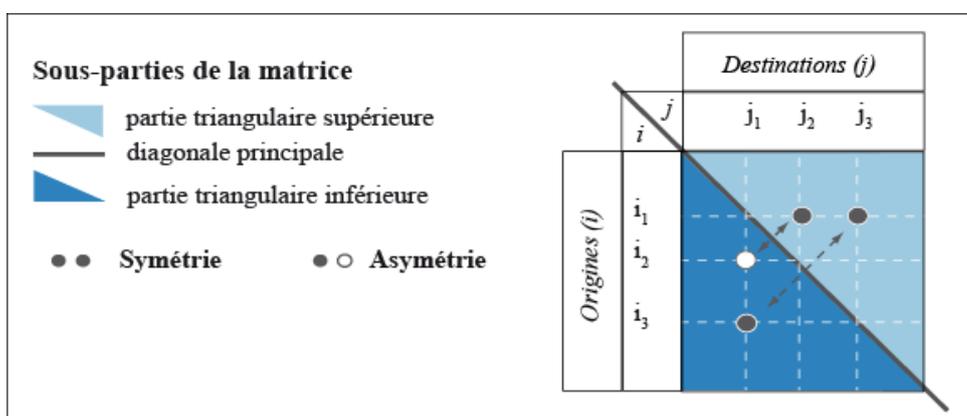
Le caractère *symétrique* d'une matrice ne s'apprécie que pour les matrices carrées où le nombre de lignes et de colonnes est identique ($N \times N$), et où les individus (i) sont les mêmes que les individus (j).

Une matrice est *symétrique* si pour tout (i, j) si ($i \neq j$), alors (F_{ij}) = (F_{ji})

Une matrice est *antisymétrique* si pour tout (i, j), si ($i \neq j$), alors (F_{ij}) = - (F_{ji}). Autrement dit (F_{ij}) et (F_{ji}) sont identiques en valeur absolue, mais de signe opposé.

Si ces deux cas de symétrie ou d'antisymétrie ne sont pas vérifiés, la matrice est *asymétrique*.

Figure 3-6. Illustration du caractère symétrique (et asymétrique) d'une matrice



Le caractère *symétrique* ou non d'une matrice n'est pas sans conséquence sur le plan cartographique, car il renseigne sur l'existence éventuelle de flux orientés, par conséquent sur la nécessité de recourir à la flèche ou non. En effet, l'existence de valeurs de flux *asymétriques* ou *antisymétriques* impose le recours à un procédé graphique orienté : la flèche ou l'une de ses variantes, ce qui n'est pas le cas des matrices *symétriques*. Autrement dit, si la matrice n'est pas symétrique, alors la flèche s'impose. Nous reviendrons dans la section 3.3 (page 127), sur les conséquences, en termes de sémiologie, de la cartographie d'une matrice de flux, selon qu'elle est symétrique ou non.

Au niveau sémiologique, la figure de la carte de flux varie, en effet, fondamentalement, en fonction du caractère symétrique ou non de la matrice. La flèche étant source de nombreuses contraintes graphiques, on voit bien l'intérêt de disposer d'une matrice de flux qui soit plutôt symétrique. Si tel n'est pas le cas et en fonction de l'objectif de la carte, il est toujours possible de la symétriser. Pour cela, plusieurs fonctions sont disponibles dans la littérature, nous en présentons les principales dans la section suivante.

3.2.2.2 La symétrisation d'une matrice

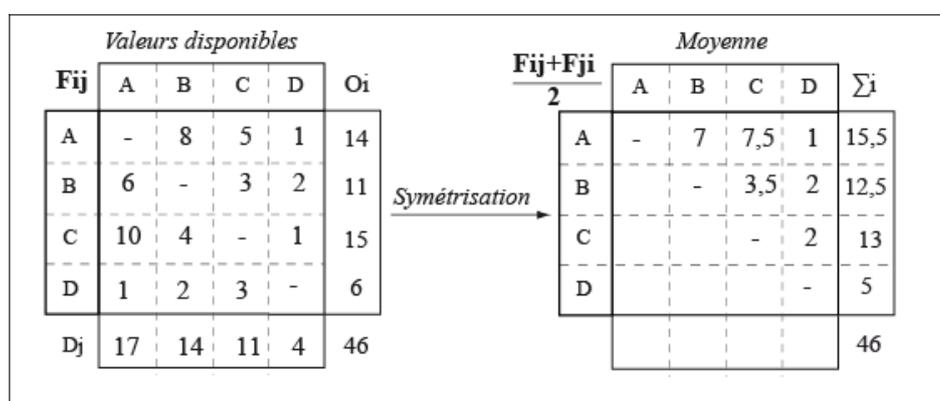
La symétrisation d'une matrice n'est envisageable que pour les matrices fermées dont le nombre de lignes est égal au nombre de colonnes. Symétriser une matrice consiste à appliquer un opérateur à la matrice complète afin de réduire sa dimension. Elle consiste à passer de ($N \times N$) cases à ($N \times N$)/2 cases renseignées. La Figure 3-7 illustre une procédure de symétrisation appliquée à notre matrice exemple. L'opération consiste ici à prendre la valeur moyenne des flux bilatéraux observés ce qui, appliquée au niveau des liens, correspond à $[(F_{ij} + F_{ji})/2]$. Nous avons réalisé la même opération pour les marges soit calculé la moyenne des flux aux lieux $(O_i + D_j)/2$.

La partie gauche de la Figure 3-7 correspond à notre matrice exemple : les valeurs de (A, B) et de (B,A) y sont différentes, respectivement égales à 8 et à 6. Sur la partie droite, la valeur (A, B) est égale à 7 [= (8+6) / 2] ; elle aurait été identique à la valeur (B, A) qui n'a pas été portée sur la figure, afin de ne pas l'alourdir par la duplication de valeurs identiques de part et d'autre de la diagonale principale. De fait, seules les cellules situées sur la partie triangulaire supérieure de la matrice sont renseignées.

Ainsi, lorsque (A, B) = (B, A), alors la matrice est *symétrique*, sinon elle ne l'est pas.

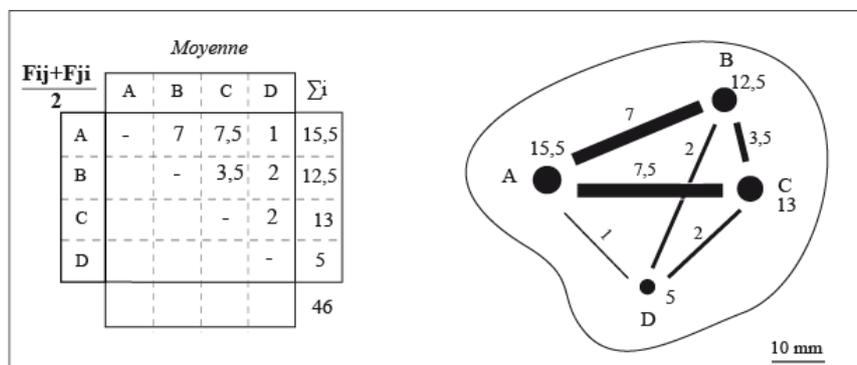
L'opération de symétrisation, quelle que soit l'opération réalisée, a pour objectif de réduire les données de la matrice, sa densité, en rendant identiques les valeurs situées de part et d'autre de la diagonale principale. Elle revient à transformer la matrice en une sous-matrice triangulaire, tout en assurant la conservation du total des flux, c'est-à-dire de l'information totale, comme illustré par la Figure 3-7 : le total général est toujours égal à 46 sur les deux matrices présentées.

Figure 3-7. Illustration d'une procédure de symétrisation d'une matrice (Fij)



Lorsque la matrice contient des valeurs qui sont symétrisées (ou symétriques), les flux dont il est question ne sont pas systématiquement orientés. C'est le cas de ceux qui sont représentés sur la Figure 3-7. C'est pourquoi certains aspects graphiques de la représentation de ces flux s'en trouvent facilités : il n'est plus nécessaire d'utiliser une flèche, et la quantité de figurés à représenter est divisée par deux, ce qui constitue un avantage dans bien des cas, surtout lorsque la matrice est dense. La Figure 3-8 illustre la cartographie d'une matrice de flux symétrique : il s'agit de la matrice présentée sur la Figure 3-7 (partie droite).

Figure 3-8. Illustration de la cartographie d'une matrice de flux symétrique

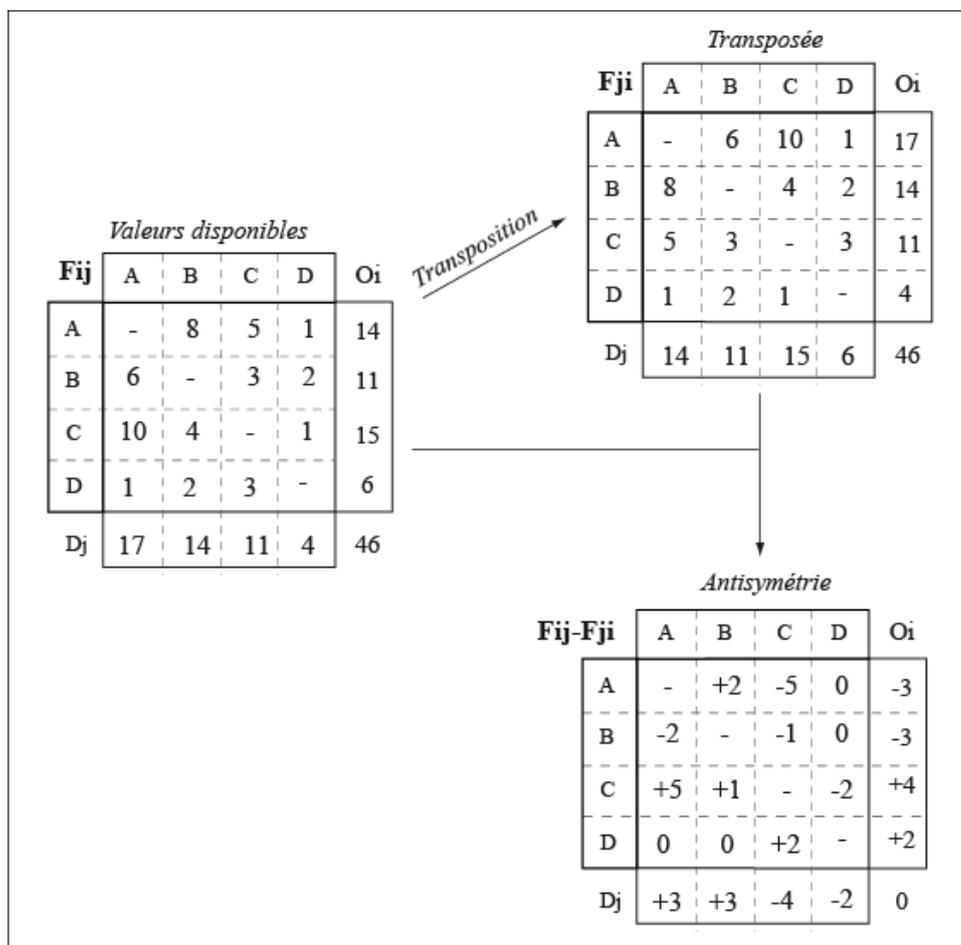


La largeur des bandes et la taille des points sont proportionnelles aux valeurs de flux correspondantes.

La symétrisation est ainsi le résultat d'une opération réalisée entre les parties de la matrice situées de part et d'autre de la diagonale principale. Elle diffère ainsi de la *transposition* ($F_{ij}^T = F_{ji}$) qui consiste simplement à permuter deux à deux les éléments situés de part et d'autre de la diagonale principale.

La Figure 3-9 présente la procédure de transposition, puis celle d'*antisymétrisation* qui consiste à retirer à (Fij) sa transposée ^T(Fij).

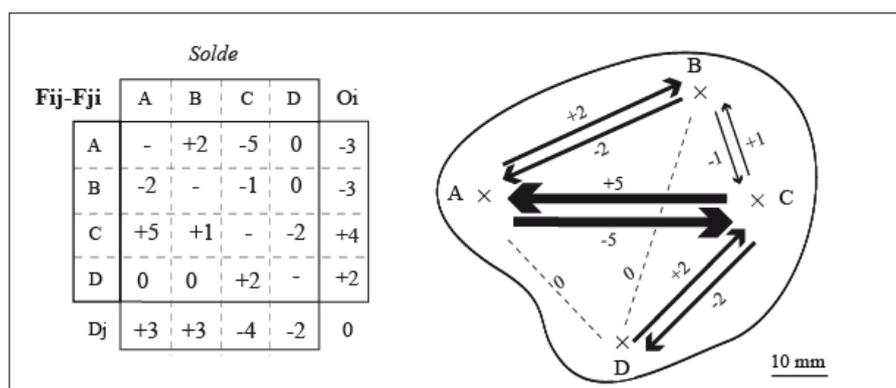
Figure 3-9. Illustration de la procédure d'antisymétrisation de la matrice (Fij)



Une matrice *antisymétrique* vérifie donc la propriété suivante :
pour tout (i, j), si (i)≠(j), alors (Fij)= - (Fji).

Lorsque la matrice est donc antisymétrique, ses valeurs sont systématiquement orientées puisqu'elles sont tantôt positives, tantôt négatives. Leur cartographie nécessite de prendre en compte ce changement de signe, ce qui conduit à une démultiplication du nombre de figurés.

La Figure 3-10 présente la cartographie directe d'une telle matrice de flux : chacune des valeurs disponibles est représentée par une flèche. On observe une démultiplication du nombre de figurés pour des valeurs de flux qui sont identiques, mais de signe opposé. Le problème qui se pose ici est lié à la nécessité de représenter l'orientation du flux, qui impose la flèche. C'est la raison pour laquelle seules les valeurs positives (ou les valeurs négatives) de ces flux sont souvent représentés par une flèche unilatérale, nous y reviendrons dans la sous-section 3.3.3.

Figure 3-10. Illustration de la cartographie d'une matrice de flux antisymétrique

La largeur des flèches est proportionnelle à la valeur du flux.

De ce point de vue, la cartographie des matrices *symétriques* apparaît plus aisée (voir Figure 3-8) puisqu'elle mobilise une bande et non une flèche, et ne nécessite pas de sélectionner les valeurs (positives ou négatives). Cependant, il convient de noter que ces gains sur une contrainte graphique ne sont pas forcément intéressants sur le plan analytique. En pratique, le fait de ne disposer que d'une matrice décomposée est parfois moins intéressant, certaines possibilités combinatoires d'analyse étant limitées.

L'exploitation du caractère symétrique d'une matrice et de ses possibilités de transposition présente un réel intérêt en cartographie des flux. Elle permet de décomposer les flux *asymétriques* tout en réduisant la taille de la matrice, ce qui revient à réduire le nombre de liens. Plusieurs approches de la décomposition des matrices peuvent être mises en œuvre, mais toutes n'intéressent pas notre cadre d'analyse cartographique. C'est pourquoi nous restreignons notre propos à l'approche de W. Tobler (1979, 1982), qui conduit à extraire deux sous-parties complémentaires de la matrice (Fij) des données initiales : il s'agit des composantes *symétrique* et *antisymétriques*, respectivement (Fij⁺) et (Fij⁻).

3.2.3 L'approche de W. Tobler : l'analyse des composantes symétriques (Fij⁺) et antisymétriques (Fij⁻) de la matrice

La proposition de W. Tobler de décomposition des matrices s'inscrit dans le cadre d'un raisonnement mené sur l'intensité des valeurs de flux et non sur les seules valeurs des flux comme précédemment. Il s'inscrit dans une démarche issue de l'algèbre des matrices, que l'auteur a donc transférées au cadre d'analyse des matrices de flux, à la fin des années 1970.

Cette proposition s'inscrit dans le cadre des méthodes qui conduisent à décomposer⁹⁷ les matrices de vecteurs. Il s'agit d'un procédé classique en algèbre linéaire, utilisé pour résoudre les systèmes d'équations linéaires présentant de 1 à n inconnues. La décomposition des matrices consiste aussi à générer n sous-ensembles réduits caractéristiques d'une matrice complète. Réduire une matrice consiste à trouver une matrice qui lui soit semblable – au sens mathématique –, tout en étant la plus simple possible.

Comme nous l'avons déjà mentionné, W. Tobler est, à l'époque, en train de résoudre un problème d'aménagement lié aux migrations. Nous détaillons sa démonstration dans l'objectif de montrer que cette solution qui apparaît générale a été découverte dans le cas particulier des migrations. L'auteur part de **l'hypothèse qu'une réduction de la matrice de départ (Fij) facilite sa représentation cartographique puisque, tout en la simplifiant, elle conserve l'information de départ**. Plusieurs

⁹⁷ Le terme mathématique exact est la réduction de l'endomorphisme, c'est-à-dire de la matrice.

méthodes de décomposition d'une matrice en sous-ensembles sont disponibles dans la littérature, mais toutes ne semblent pas, *a priori*, transférables en l'état au cas des matrices de flux.

W. Tobler démontre que la meilleure façon de cartographier des flux est de représenter des champs de vecteurs⁹⁸ calculés sur ce qu'il nomme « *la composante négative des flux* » notée (Fij⁻) et inversement, sur « *la composante positive des flux* » notée (Fij⁺). Les matrices de flux asymétriques (Fij) correspondent, en effet, à « [...] ces tables [qui] peuvent être décomposées en deux parties, une partie symétrique et une partie antisymétrique. Pour les statisticiens de l'auditoire la variance totale peut également être divisée en ces deux parties » (Tobler, 1979a, 1982), de la manière suivante :

$$(Fij) = [(Fij^+) + (Fij^-)] / 2$$

$$(Fij^+) = (Fij + Fji) \text{ et } (Fij^-) = (Fji - Fij),$$

$$\text{avec } (Fij^-) = - (Fij^+)$$

Concernant la formulation, (Fij⁺) et (Fij⁻) correspondent à deux composantes élémentaires de (Fij), issues d'une procédure de symétrisation mais pas seulement : elles sont aussi complémentaires. Cela étant, pour faciliter la lecture de la suite, nous allons considérer que la carte du (Fij⁺) correspond à celle du **volume bilatéral du flux** et celle du (Fij⁻) au **transfert net des échanges, au solde bilatéral** (à la balance de J. Bertin déjà évoquée), sans doublement des figurés. Il est également intéressant de noter l'analogie de cette proposition de W. Tobler avec celles D. Courgeau (1980), présentée dans L'Encart 3-1.

Encart 3-1. Sur le calcul de la migration totale et de la migration nette

D. Courgeau présente la « *migration totale* » comme la somme des flux bilatéraux (Fij + Fji) qu'il convient de ne pas confondre avec le « *total des migrations* » qui correspond à la somme de l'ensemble des valeurs de la matrice $\Sigma(Fij)$. L'auteur définit aussi la « *migration nette* » (Courgeau, 1980 : 84) en tant que différence entre les flux bilatéraux (Fji - Fij) et précise qu'elle ne s'applique qu'aux individus, c'est-à-dire qu'aux matrices dont les valeurs décrivent des effectifs de migrants. La *migration nette* ne s'appliquerait pas aux déplacements de ces mêmes migrants, aux migrations. L'auteur précise qu'une zone ayant une « *migration nette* » nulle peut toutefois avoir connu une variation interne forte de « certaines catégories » de populations qui la composent, sans pour autant que cela soit visible. De la même façon, une zone qui présente une « *migration nette* » fortement positive ou négative peut conserver une structure interne identique sans que l'effectif global de migrants soit modifié.

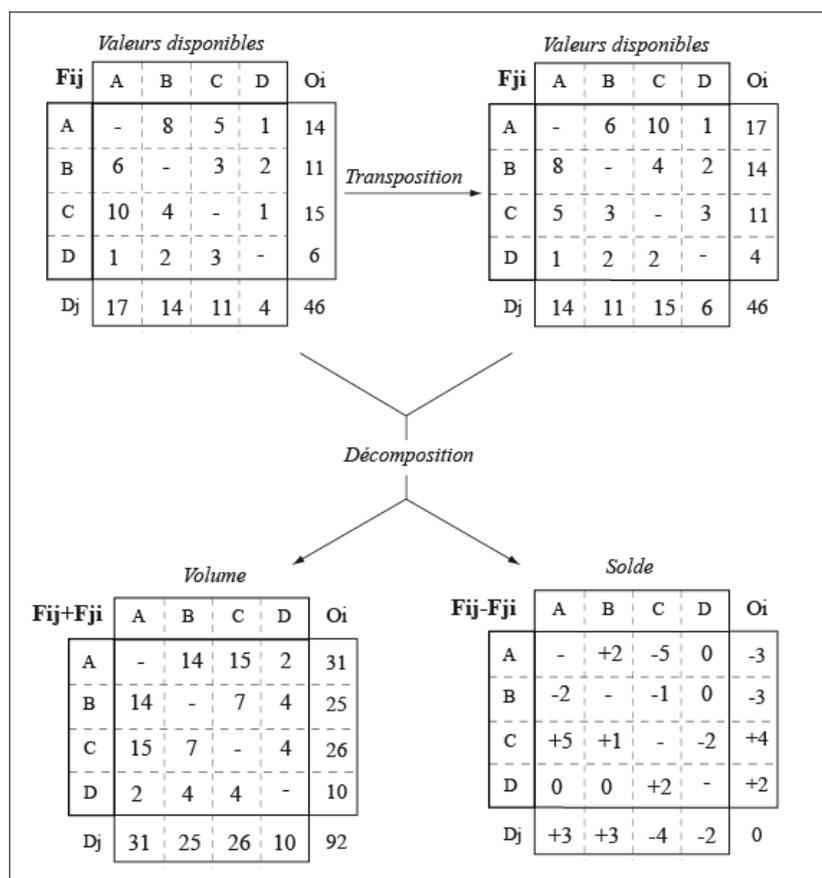
Il est dès lors intéressant de conduire une « *analyse différenciée des flux* », ce qui suppose d'articuler différentes matrices d'une même zone d'étude décomposée en plusieurs catégories d'individus. Il convient de noter que la « *migration nette* », fondée sur l'additivité des valeurs de flux bilatérales, diffère du « *nombre net de migrants* » où l'additivité ne fonctionne pas puisque nous avons vu dans le chapitre 1, qu'un migrant pouvait avoir effectué plusieurs migrations entre deux zones, sur une période, et non l'inverse. L'auteur suggère également la définition de « *taux d'émigrations* » (ou d'immigrations) bruts ou nets qui se rapportent à la population moyenne de la zone, pendant la période considérée. Aussi le recours à deux indices : « *l'indice d'efficacité* » et « *l'indice de compensation* » (Courgeau, 1980 : 63). L'indice d'efficacité des zones dans le système d'échanges résulte de la part de la « *migration nette* » dans le « *total des flux* » d'une zone tandis que « *l'indice de compensation* » correspond à une forme **d'asymétrie** des échanges : il traduit le rapport des « *migrations nettes* » sur les « *migrations totales* » (Courgeau, 1980 : 141).

⁹⁸ Ces champs de vecteurs devant le conduire à la notion d'attractivité.

La décomposition des matrices de flux en deux composantes, telle quelle est proposée par W. Tobler, présente l'avantage d'améliorer la compréhension des processus sous-jacents au système des flux. D'après Cauvin et Reymond (1991), la cartographie du (F_{ij}^+) exprime le « *fond vital du système* », sa stabilité tandis que celle du (F_{ij}) exprime la « *direction nette des échanges* » (Cauvin C., Reymond H, 1991 : 468). Elle présente aussi l'avantage de symétriser la matrice (F_{ij}) , ce qui conduit à une réduction quantitative du nombre de figurés à représenter dans les deux cas.

La Figure 3-11 illustre la décomposition de la matrice initiale (F_{ij}) en deux composantes, qui place l'analyste devant quatre matrices issues des données (F_{ij}) initiales.

Figure 3-11. Illustration de la décomposition de la matrice (F_{ij}) en composantes (F_{ij}^+) et (F_{ij}^-)



D'après la Figure 3-11, **on dispose de quatre matrices de flux issues des données observées (F_{ij}) , pour une analyse cartographique des flux.** Elles sont décrites de la manière suivante.

- La première correspond aux matrices de flux observés (F_{ij}) qui sont généralement *asymétriques*, ce qui signifie que pour tout (i) différent de (j) , (F_{ij}) n'est pas nécessairement différent de (F_{ji}) .
- La seconde matrice (F_{ji}) est obtenue par la transposition de celle des données initiales, elle n'est possible que si la matrice est fermée et carrée, si les (i) lignes sont identiques aux (j) colonnes.
- La troisième matrice est celle des matrices *symétriques*, positives ou négatives, où pour tout (i) différent de (j) , $(F_{ij}) = (F_{ji})$. Ces matrices symétriques présentent des valeurs situées de part et d'autre de la diagonale principale qui sont identiques.
- La quatrième matrice concerne les matrices *antisymétriques*. Elles présentent des valeurs qui sont égales en valeur absolue de part et d'autre de la diagonale principale mais de signe opposé. Ainsi, pour tout $(i) \neq (j)$, $(F_{ij}) = - (F_{ji})$.

Ces différentes matrices principales sont obtenues en appliquant le raisonnement proposé par W. Tobler, qui permet de représenter la partie stable ou la partie dynamique du système de flux.

Dans le cadre classique d'un raisonnement scientifique, la posture de W. Tobler n'est pas réductrice car elle procure un cadre général d'analyse qui nous permet de présenter une typologie des méthodes de cartographie de flux. En effet, quel que soit le raisonnement adopté et la méthode (descriptive, explicative, prédictive) ayant permis de générer la matrice de flux à cartographier, l'analyste se retrouve toujours devant l'une ou l'autre des matrices de la Figure 3-11. Autrement dit, **quel que soit le cas, la matrice dont on dispose est toujours soit symétrique, soit asymétrique, soit antisymétrique.**

En conséquence, dans la section suivante, nous poursuivons dans la logique de W. Tobler et présentons les méthodes de cartographie spécifiques à chacune des quatre matrices de la Figure 3-11. Ces matrices correspondent en réalité à trois familles de méthodes et non pas quatre, puisque les méthodes de cartographie applicables à une matrice (F_{ij}) et à sa transposée (F_{ji}) sont rigoureusement identiques. Une quatrième famille de méthodes peut toutefois être introduite pour les matrices complexes, dont l'élaboration combine l'une ou l'autre des trois matrices de flux principales. Cependant, elle est à manier avec précaution car la formulation générale de *l'asymétrie*, résultant du rapport du solde sur le volume, entraîne la perte de la propriété générale d'additivité en lignes et en colonnes de la matrice.

3.3 Les trois grandes méthodes de cartographie des matrices de flux

Nous proposons dans cette section une typologie des méthodes classiques de cartographie de flux, en trois grandes familles : chacune d'elle s'appliquant à un type particulier de matrice de flux. D'après la Figure 3-11, quel que soit le raisonnement qui a permis de les générer, les matrices de flux se distinguent en fonction de leur caractère symétrique ou non par rapport à la diagonale principale, c'est-à-dire selon qu'elles sont *asymétriques* (F_{ij}), *symétriques* ($F_{ij}=F_{ji}$) ou *antisymétriques* ($F_{ij}=-F_{ji}$).

A chacune de ces trois familles de matrices correspond, sur le plan sémiologique, une méthode principale de cartographie. Cela porte donc à trois le nombre de figurés distincts et non directement substituables permettant de représenter convenablement les flux – toujours du point de vue des liens – de chacune de ces matrices. Ainsi, nous pouvons affirmer que :

- **le figuré usuel des cartes de flux asymétriques est la flèche bilatérale** (ou la double flèche) ;
- **le figuré des matrices symétriques est la bande** (ou le lien de taille variable) ;
- **le figuré adéquat pour cartographier les matrices antisymétriques est la flèche unilatérale.**

Pour illustrer l'application de ces trois méthodes de sémiologie de la carte de flux, nous avons recours à notre matrice (F_{ij}) exemple. Celle-ci étant fictive, nous illustrerons leur application sur la matrice des flux interdistricts de l'ex-Tchécoslovaquie (1990).

La cartographie des matrices de la Figure 3-11 nécessite de placer sur la carte, 6 figurés qui correspondent aux flux observés entre les quatre lieux et quatre autres figurés pour représenter l'effet des flux sur la distribution des lieux. Nous excluons ici les méthodes de la cartographie statistique choroplèthe s'appliquant aux cas des flux analysés du point de vue des lieux.

Pour cartographier ces matrices du point de vue des liens, nous faisons le choix d'utiliser un trait droit, un lien rectilinéaire pour illustrer le flux reliant une origine à une destination. Cela ne pose pas de problèmes particuliers pour les raisons suivantes : ce lien est censé illustrer un trajet direct (origine-destination), souvent mesuré dans la métrique euclidienne du plus court chemin. Le tracé droit correspond donc, de façon conventionnelle, à la distance mesurée à vol d'oiseau que celui-ci décrit. Par ailleurs, le cheminement exact du flux entre deux lieux n'étant pas connu, le fait que ce trait soit

droit ou non, n'a pas de réelle importance pour la compréhension. Son intérêt est surtout esthétique. Nous faisons donc le choix du trait, car il n'a pas besoin d'être incurvé. Cela signifie que nous ne lui appliquons aucune transformation d'ordre graphique ou même cartographique. Ce trait droit, lorsqu'il est orienté sera, en revanche, éventuellement assorti d'un sens, en fonction de la polarité de la valeur de flux de la matrice correspondante. Autrement dit, lorsqu'il est nécessaire de représenter le sens du flux, une pointe sera apposée au trait, figurant ainsi une flèche (ou tout autre procédé permettant de représenter une orientation). Pour représenter les flux observés du point de vue des lieux, nous faisons le choix du point.

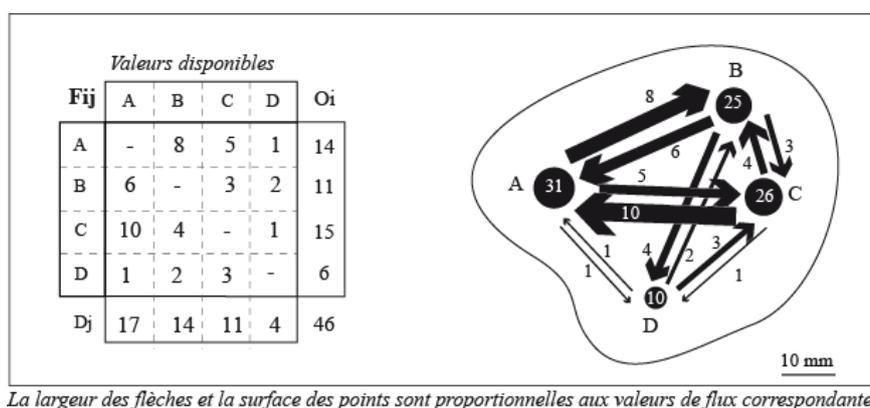
Dans les quatre sous-sections suivantes, nous présentons les aspects de sémiologie graphique relatifs à la cartographie des principales matrices de flux, présentées sur la Figure 3-11.

3.3.1 La cartographie des matrices de flux asymétriques : l'exemple des valeurs observées (Fij)

La première famille de méthode est celle des matrices de flux *asymétriques*, telles que (Fij), qui décrivent le transfert d'une quantité positive entre deux lieux. Cette matrice correspond au cas le plus simple, souvent retrouvé en pratique, où l'on cartographie les valeurs de flux observées à l'aide d'une flèche (ou de tout autre figuré orienté, le raisonnement étant le même).

Sur le plan sémiologique, les flux asymétriques sont en général représentés par des flèches, puisqu'ils sont orientés (voir Figure 3-12) ; la flèche indique toujours la quantité et le sens du flux. On peut représenter le volume de flux au niveau des lieux par une implantation ponctuelle.

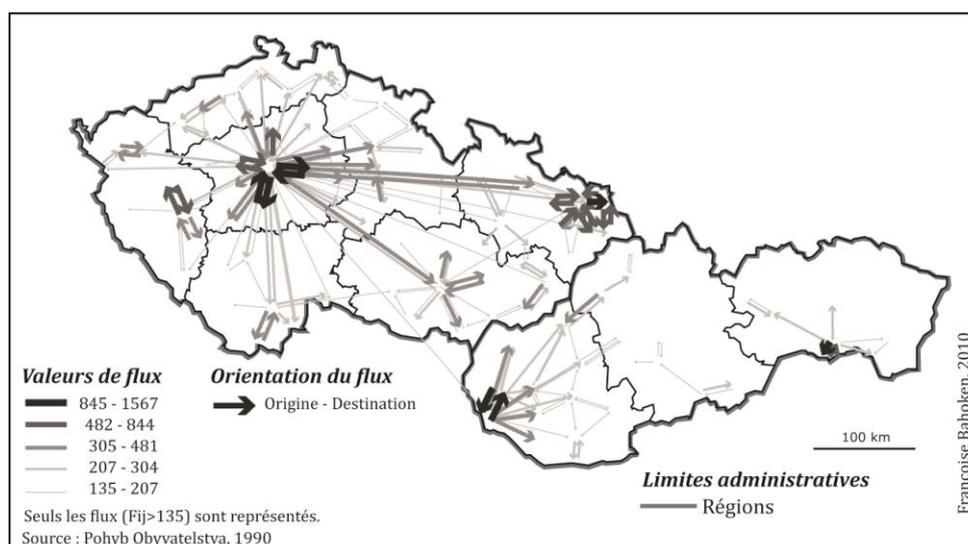
Figure 3-12. Illustration de la cartographie classique (liens, lieux) de la matrice (Fij)



La cartographie des flux (Fij) de l'ex-Tchécoslovaquie est présentée sur la Figure 3-13. Les valeurs de flux ont fait l'objet d'une sélection pour garantir la lisibilité de la carte, puisque nous avons vu (chapitre 2) que cette matrice était soumise fortement à l'*effet-spaghetti* (voir Figure 2.19, page 85). Seuls 2% des liens sont présents sur la Figure 3-13.

Les valeurs de flux concernées ont aussi été regroupées en cinq classes de valeurs afin de réduire l'hétérogénéité de taille des liens. A noter que d'un point de vue sémiologique, l'ajout de la *Valeur* est inutile, voire incorrecte. Elle peut toutefois être justifiée par des raisons esthétiques ou thématiques visant, par exemple, à renforcer l'attention sur les plus grosses valeurs. Ainsi, à chacune des classes de valeur correspond une classe de taille des figurés. Enfin, à chacune de ces cinq classes correspond – de façon facultative – une nuance de teinte qui correspond ici à une intensité de niveau de gris définie à l'aide de l'outil *Color Brewer*.

Figure 3-13. Cartographie des flux de l'ex-Tchécoslovaquie (1990)



L'attention est attirée sur le fait que les questions relatives aux choix des différents critères (d'esthétique, de sélection et de segmentation de la série) seront examinées en seconde partie (chapitre 5, 8 et 9).

Sur le plan théorique, le problème des cartes réalisées à partir des données initiales (F_{ij}) est qu'elles comparent la valeur de l'ensemble de l'information, ce qui peut constituer une gêne pour la compréhension fine de la structure spatiale des échanges. Il s'agit de la manifestation d'un *effet d'amalgame*⁹⁹ résultant du mélange d'informations de nature différente, dont nous verrons qu'elles portent sur les aspects dynamiques et non des flux. Cet effet résultant d'une exploitation de l'ensemble de la matrice (F_{ij}), nous proposons de la décomposer en (F_{ij}^+) et (F_{ij}^-).

La cartographie sur différentes cartes, de chacune de ces parties complémentaires constitue, en effet, une amélioration qualitative intéressante du cas classique de cartographie de la matrice (F_{ij}). La section suivante présente la seconde famille méthode de cartographie des matrices de flux, elle s'applique au cas des matrices *symétriques* (F_{ij}^+).

3.3.2 La cartographie des matrices de flux symétriques, l'exemple des volumes de flux bilatéraux (F_{ij}^+)

La seconde famille de méthode est celle des matrices *symétriques* ($F_{ij}=F_{ji}$), elle est illustrée sur la partie basse de la Figure 3-14. La carte correspondante est celle des *volume-volume* qui décrit le *volume de flux bilatéral* ($F_{ij}+F_{ji}$) et celui de leur effet aux lieux (O_i+D_j).

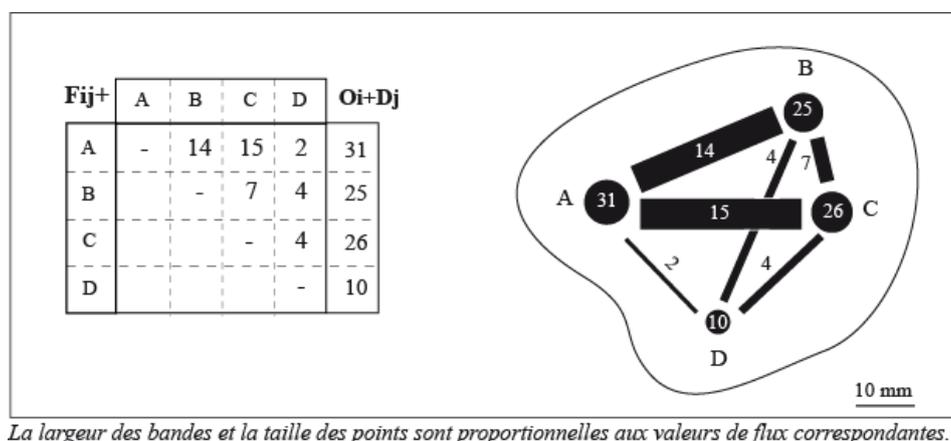
La cartographie naturelle du *volume bilatéral* est formée (sur la partie supérieure de la Figure 3-14) par un ensemble de flèches bilatérales qui illustrent la même information et sont de ce fait identiques. Elles sont donc redondantes puisque deux flèches de même valeur relient le même couple de lieux. Etant donné que la même quantité d'information est échangée entre les couples de lieux, il n'est pas nécessaire de conserver la double-flèche, ni d'ailleurs la flèche pour réaliser une carte du volume de flux et, plus généralement, pour cartographier les matrices de flux *symétriques* : une bande suffit.

C'est pourquoi la cartographie des matrices symétriques ne nécessite pas le recours à la flèche, la bande proportionnelle permettant de représenter efficacement la quantité transférée entre deux lieux, comme illustré par la partie basse de la Figure 3-14. Il n'est par conséquent pas nécessaire de conserver cette double-flèche ni d'ailleurs la flèche. Le figuré qui convient à la représentation des

⁹⁹ Cette définition de l'*effet d'amalgame* sera étendue dans la section 5.2 (page 216).

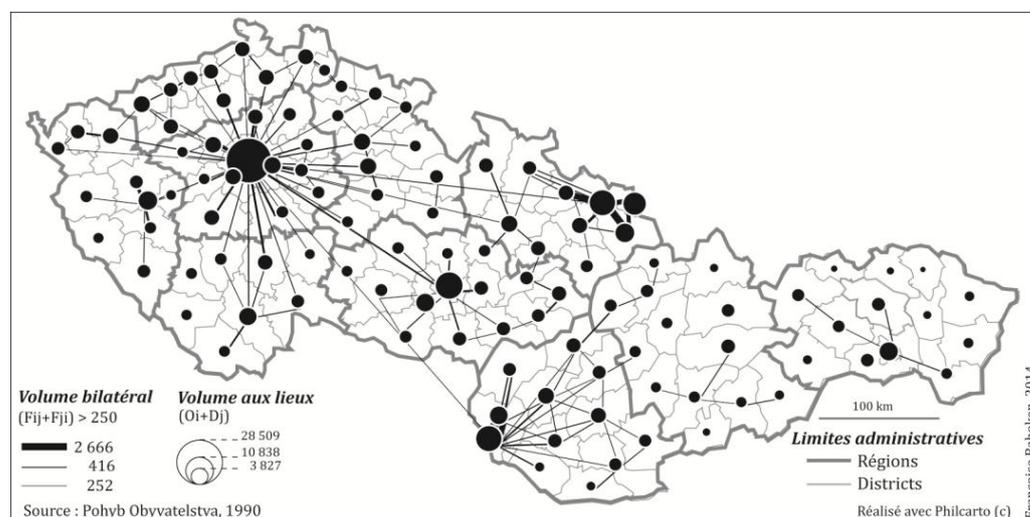
volume-volume, à la cartographie des matrices symétriques, est la bande proportionnelle, efficace pour décrire la quantité transférée entre deux lieux.

Figure 3-14. Cartographie de la composante symétrique (F_{ij}^+) des flux (version 1)



Sur une carte des *volumes de flux*, ce que l'observateur est censé percevoir instantanément, c'est bien une variation de la taille des *largeurs* de bandes qui correspond aux différentes valeurs de flux, pas celle de leur *longueur* puisque le trait est droit, ni de leur *hauteur*, la figure étant réalisée en deux dimensions. L'application de la *Taille*, en cartographie de flux, correspondant bien à une variation de la *largeur* des bandes. La mise en œuvre de leur surface (issue du produit *largeur*longueur*) proportionnellement à leur valeur, est intenable en raison de la contrainte spatiale de position des lieux. La Figure 3-15 présente la carte des volumes de flux de l'ex-Tchécoslovaquie.

Figure 3-15. Cartographie du volume bilatéral (F_{ij}^+) de flux de l'ex-Tchécoslovaquie (1990)



La carte de la Figure 3-15 illustre les représentations complémentaires (*liens-lieux*) des *volumes* (bilatéral entre les lieux et aux lieux), par la représentation proportionnelle des figurés correspondants : la bande et le point. Cette représentation donne une bonne idée de la structure des échanges et de l'importance des lieux : leur hiérarchie est perceptible. Sur le plan sémiologique et plus généralement méthodologique, l'intérêt de cette représentation est l'économie liée à l'absence de flèches qui illustrent l'orientation du déplacement, ainsi qu'au volume de données moins important à manipuler puisque la matrice symétrique. Le procédé ne convient toutefois qu'aux matrices *symétriques*. Dans le

cas contraire, souvent rencontré en pratique, qui concerne les matrices *antisymétriques*, il faut utiliser la méthode du *bilan net*, en réalisant une carte de flux de type *solde-solde*.

3.3.3 La cartographie des matrices de flux antisymétriques : l'exemple des transferts nets bilatéraux (F_{ij}^-)

La troisième famille de méthode est celle des matrices antisymétriques, que nous illustrons avec la matrice (F_{ij}).

Les données de la matrice (F_{ij}^-) sont plus difficiles à manier que celles du (F_{ij}^+), car le sens du flux a une signification. La valeur de (F_{ij}) étant potentiellement différente de celle de (F_{ji}), celle de (F_{ij}) est susceptible de varier en fonction de la formulation de calcul, comme illustré par la Figure 3-16. En fait, l'une des matrices de cette figure étant la transposée de l'autre : $(F_{ij}-F_{ji}) = {}^T(F_{ij}-F_{ji}) = (F_{ji}-F_{ij})$

Figure 3-16. Illustration de la variation de la valeur du solde, en fonction de la formule

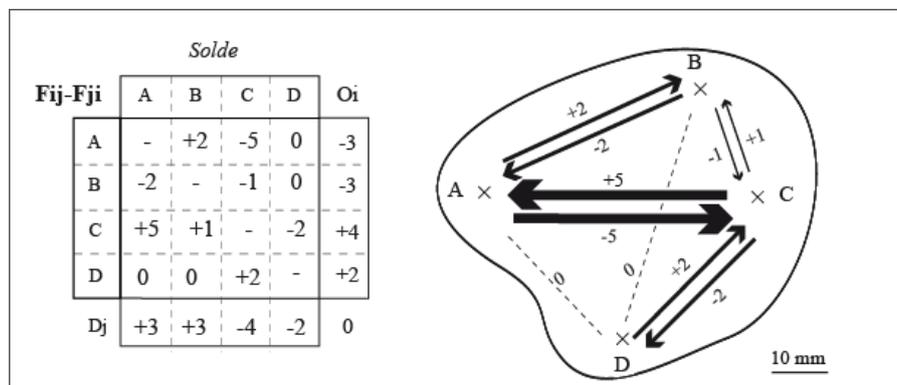
Solde						Solde					
$F_{ij}-F_{ji}$	A	B	C	D	O _i	$F_{ji}-F_{ij}$	A	B	C	D	O _i
A	-	+2	-5	0	-3	A	-	-2	+5	0	+3
B	-2	-	-1	0	-3	B	+2	-	+1	0	+3
C	+5	+1	-	-2	+4	C	-5	-1	-	+2	-4
D	0	0	+2	-	+2	D	0	0	-2	-	-2
D _j	+3	+3	-4	-2	0	D _j	-3	-3	+4	+2	0

La formule de calcul du *solde bilatéral* n'est pas sans importance, car elle est susceptible de modifier les résultats de l'analyse et l'apparence de la carte, puisque les flèches n'auront pas le même sens en fonction des cas et cela n'est pas négligeable lorsque les lieux d'origine ne correspondent pas à ceux de destination. C'est le cas, par exemple, des matrices de flux commerciaux qui représentent l'exportation (F_{ij}) ou l'importation (F_{ji}) quand celle des flux de navetteurs quotidiens décrivent l'attractivité des bassins d'emplois ou l'émissivité des zones résidentielles plus ou moins bien dotées en emplois. Elle nécessite d'être appliquée à la thématique analysée. Pour notre part, nous faisons le choix de la formulation $(F_{ij}-F_{ji})=S_{ij}$ qui correspond au solde bilatéral ou transfert net que nous avons évoqués dans la section 1.1.1 du chapitre 1 (page 31).

Sur le plan sémiologique, la cartographie du *transfert net* diffère fondamentalement de celle du *volume bilatéral*. En effet, pour tout $(i) \neq (j)$, (F_{ij}^-) n'est pas différent de (F_{ji}^-) mais est son exact opposé : $(F_{ij}) = -(F_{ji})$, sa transposée, tandis que dans la matrice (F_{ij}^+), pour tout $(i) \neq (j)$, (F_{ij}) est égal à (F_{ji}). Cette spécificité de la matrice (F_{ji}^-) nécessite de prendre en compte le sens du flux dans sa cartographie, en utilisant une flèche, comme dans le cas de la représentation de la matrice (F_{ij}) illustré par la Figure 3-12. Cependant, à la différence de cette matrice (F_{ij}) où les valeurs (F_{ij}) et (F_{ji}) qu'elle contient renseignent sur des informations différentes ($F_{ij} \neq F_{ji}$), les valeurs de la matrice (F_{ij}^-) présentent la même information de part et d'autre de la diagonale principale mais de signe opposé, sauf lorsqu'elle est nulle ($F_{ij}=0$), que le bilan est équilibré. Du point de vue des lieux, cette différence entre les valeurs de flux entrants et sortants correspond à la variation de leur population, tandis que du point de vue des liens, elle décrit le *transfert net* dont il faut représenter le sens, par une flèche, pour distinguer les lieux qui perdent de ceux qui gagnent.

La cartographie du **transfert net** (F_{ij}) consiste à représenter l'ensemble des valeurs de la matrice sur la même carte, à l'instar de la cartographie de la matrice (F_{ij}). Déjà présentée sur la Figure 3-11, nous la reprenons ici pour le besoin de la démonstration.

Figure 3-17. Illustration de la cartographie classique du transfert net (F_{ij})



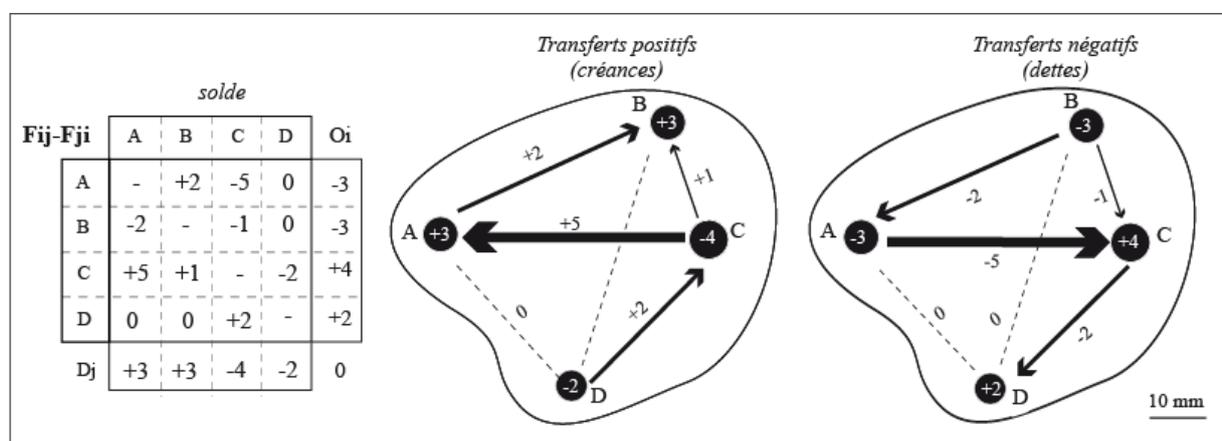
La largeur des flèches est proportionnelle à la valeur du flux.

La représentation sur la même carte de valeurs de flux identiques, mais symétriquement négatives et positives entraîne une confusion dans la caractérisation des lieux émetteurs et récepteurs et une démultiplication des figurés de tailles identiques et d'orientations opposées. Nous l'avons déjà mentionné et poursuivons le raisonnement.

En effet, sur la Figure 3-17, lorsque (F_{ij}) vaut +2, cela signifie que (j) a reçu 2 pièces de (i), soit (+2) qu'il n'en a donné [$(F_{ij})=2$] ; ou bien que (i) a donné 2 pièces à (j) soit (-2) de plus qu'il n'en a reçu [$(F_{ji})=-2$]. De ce point de vue, la Figure 3-17 réalisée sur l'ensemble des valeurs de la matrice (F_{ij}) représente à la fois les gains et les pertes, respectivement les créances et les dettes si l'on file cette métaphore financière.

Pour remédier à cet écueil lié à la prise en compte du caractère **antisymétrique** des flux, conduisant à représenter sur la même carte les créances et les dettes, **la pratique usuelle consiste à ne réaliser qu'une seule des deux cartes : l'une pour chacune des situations**. La Figure 3-18 présente ces deux cartes de flux, aux valeurs respectives strictement positives (les créances) ou négatives (les dettes) issues d'une sélection de la matrice (F_{ij}).

Figure 3-18. Illustration de la cartographie du transfert net (F_{ij}) (variante en deux cartes)



La largeur des flèches et la taille des points sont proportionnelle à la valeur du flux correspondante.

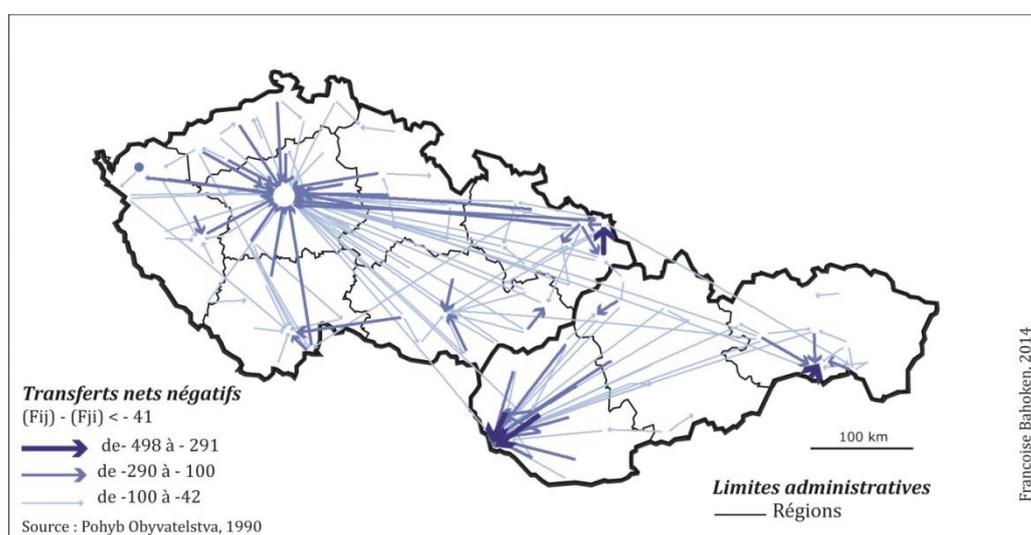
Cette méthode présente aussi l'avantage d'autoriser la représentation d'une somme marginale. Ici, **le figuré ponctuel symbolise la somme marginale dirigée dans le sens du flux entrant dans chaque lieu**. Ainsi, sur la carte des transferts positifs : le lieu A reçoit +5 de C et émet +2 vers B, son bilan est donc de +3.

Les cartes du *transfert net*, positif et négatif, non symétrisées, présentent la particularité d'être identiques, à la différence près du sens des flèches. Il est donc inutile de réaliser les deux cartes, le choix de l'une ou de l'autre dépendant de la thématique, puisqu'elle entraîne un changement de l'orientation des flèches.

Pour illustrer la cartographie du *transfert net* appliquée aux flux de l'ex-Tchécoslovaquie (1990), nous présentons sur la Figure 3-19 une carte des seules valeurs négatives du *transfert net* ($F_{ij} < 0$). Il est important de noter que dans le cas où les lieux d'origine diffèrent de ceux de destination (importations / exportations ou domicile / travail, par exemple), le choix de représenter les valeurs négatives ou positives est crucial sur le plan thématique. Ce n'est pas le cas ici, car la matrice de l'ex-Tchécoslovaquie décrit une migration résidentielle (lieu de résidence * lieu de résidence).

La Figure 3-19 présente l'approche usuelle réalisée avec des seules flèches : le bilan des flux n'y est pas représenté du point de vue des lieux (c'est-à-dire le *bilan net*), mais seulement au niveau des liens. Ces derniers sont valués et orientés, afin d'indiquer le sens du transfert et son importance. Le choix de la teinte bleue est ici purement arbitraire.

Figure 3-19. Cartographie du *transfert net* ($F_{ij} < 0$) des flux de l'ex-Tchécoslovaquie (1990)

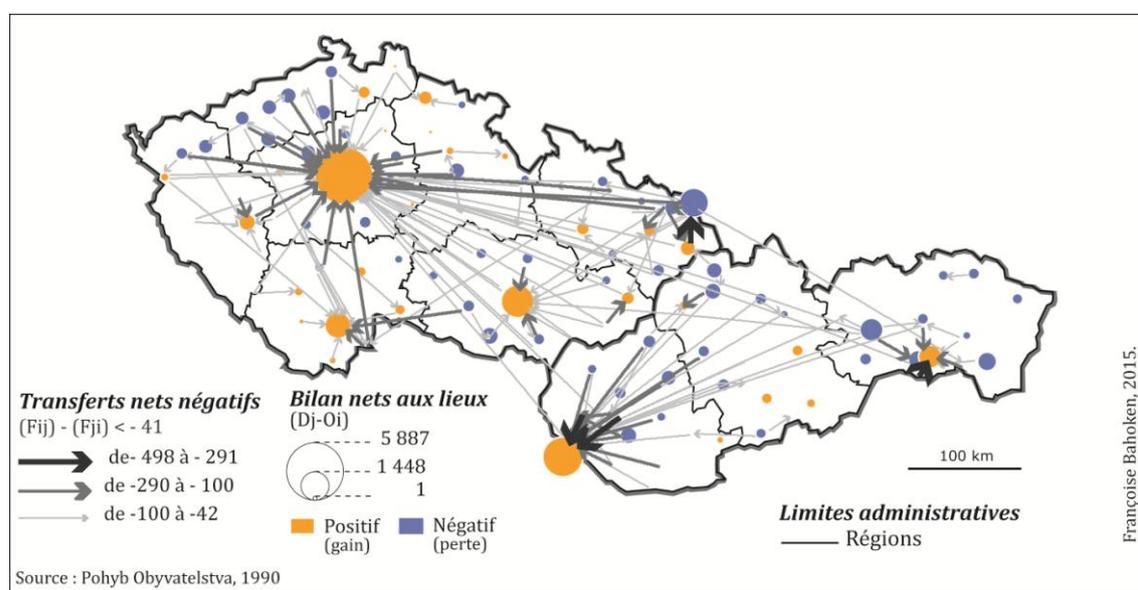


De manière générale, le choix de cartographier les valeurs négatives ou positives tient à la problématique sous-jacente. Il ne pose pas de problème particulier. En revanche, ce qui en pose un, c'est la proposition de W. Tobler (n.d-b) de représenter la valeur absolue de (F_{ij}). Le problème est en effet *a priori* insolvable en raison de l'orientation du flux. Aussi, plutôt que d'investiguer dans cette direction, nous préférons ajouter à cette figure la représentation du *bilan net* en introduisant deux composantes de sémiologie qui permettront de distinguer les lieux gagnants des perdants.

Ainsi, du point de vue des lieux, les symboles ajoutés sont des cercles dont la surface est proportionnelle à la valeur du bilan. De ce fait, leur *Taille* est équivalente à la somme des *largeurs* des bandes entrantes (ou sortantes) de chacun des lieux. Ces symboles seront également assortis d'une coloration discrétisée en fonction de la polarité du bilan (positive / négative) de sorte que les gagnants (que nous cartographions en jaune) se différencient clairement des perdants (que nous conservons en bleu). Nous conservons les têtes de flèches – éludant donc la suggestion de W. Tobler – afin de ne pas

généraliser d'ambiguïtés. En effet, il n'est apparemment pas possible de connaître l'orientation des flux par la seule connaissance de l'information du bilan aux lieux. C'est pourquoi la Figure 3-20 illustre cette proposition, pour les flux de l'ex-Tchécoslovaquie (1990), en reprenant les figurés linéaires de la Figure 3-19.

Figure 3-20. Proposition de cartographie du *transfert net* des flux de l'ex-Tchécoslovaquie (1990) : $(F_{ij}) > 0$



L'intérêt de cette proposition de cartographie du (F_{ij}) est bien de ne réaliser qu'une seule carte qui combine en outre la représentation du *transfert net* et du *bilan net*. En effet, pour deux lieux quelconques, on voit bien sur la Figure 3-20 que les valeurs de flux fortes (illustrées par de gros liens) impliquent des lieux qui gagnent ou qui perdent fortement soit parce qu'ils reçoivent de nombreux petits flux, soit parce qu'ils reçoivent peu de flux mais qui présentent des valeurs importantes : la différence entre ces deux éventualités dépend de la *longueur* des liens, c'est-à-dire d'une acception de la distance parcourue par le flux.

La proposition de W. Tobler de décomposer la matrice initiale (F_{ij}) a donc une conséquence forte, puisqu'elle conduit *in fine* à réaliser deux cartes à partir d'une même matrice de données observées, chacune d'elle illustrant une partie de l'information disponible dans la matrice complète : l'une pour la composante symétrique (F_{ij}^+) et l'autre pour la composante antisymétrique (F_{ij}^-) . Si la cartographie du (F_{ij}^+) ne pose pas de problèmes particuliers, celle du (F_{ij}^-) a pu être controversée et cela, outre le fait qu'il faille choisir ou non de réaliser l'une et/ou l'autre des deux cartes correspondantes. C'est pourquoi certains auteurs ont suggéré de cartographier plutôt un indicateur de l'*asymétrie* des flux, correspondant à un autre compromis de représentation qui associe le *volume bilatéral* et le *transfert net* sur la même carte.

L'explication avancée par les détracteurs du solde est avant tout théorique¹⁰⁰ et ne conduit pas à une nouvelle famille de méthode de cartographie. Sa principale caractéristique est de rendre plus complexe

¹⁰⁰ A. Rogers (1990), dans un texte qui prend la forme d'un plaidoyer pour la notion de migrant net, souligne la déficience de la migration nette pour analyser des mobilités. L'auteur liste pour cela différents problèmes qui découlent de son utilisation, parmi lesquels le fait que la représentation des transferts nets masque les gros flux en valeur absolue. Son utilisation serait alors préjudiciable à la modélisation des flux, puisque ses fondements théoriques sont inadaptes. C'est, en effet, comme s'il existait une corrélation entre le flux entrant et le flux sortant de chaque zone (Rogers, 1990 : 283, 299) ce qui n'est pas nécessairement le cas. L'auteur suggère alors soit de représenter le volume bilatéral (Rogers, 1990 : 299), soit un indicateur de l'asymétrie des flux.

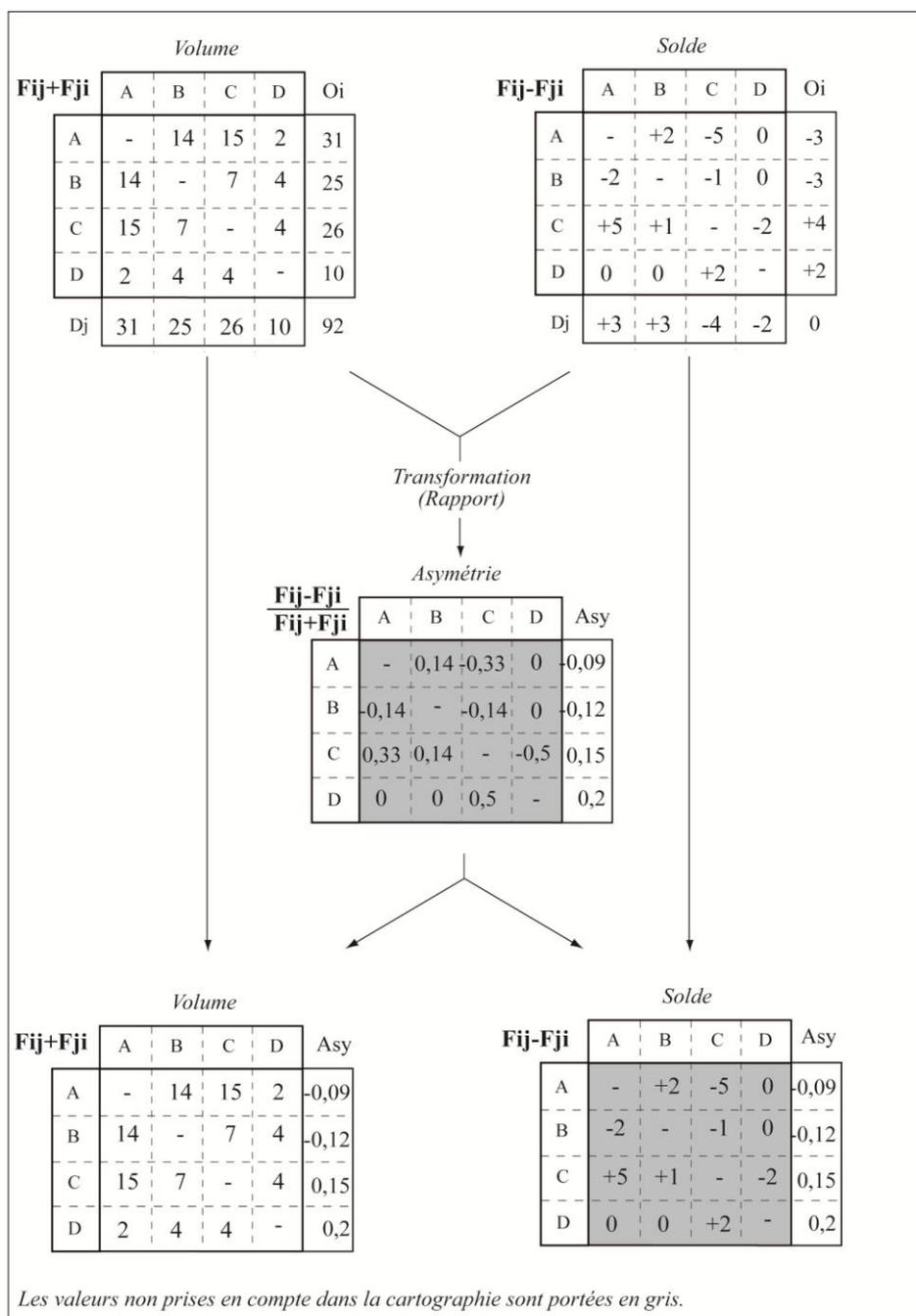
les trois principales familles, en combinant les aspects de sémiologie. Nous présentons dans la section suivante un exemple de carte de flux correspondant à la représentation d'un indice d'asymétrie des flux, traduisant l'attractivité des lieux, en utilisant la méthode de cartographie de type *volume-volume-asymétrie*.

3.3.4 La cartographie des matrices de flux issues d'une combinaison entre (F_{ij}^+) et (F_{ij}^-) , l'exemple de l'asymétrie (W_{ij})

La quatrième famille de méthode est celle des matrices de flux issues d'un compromis de représentation entre (F_{ij}^+) et (F_{ij}^-) . Nous l'illustrons à l'aide de *l'asymétrie* (W_{ij}) , formulée comme le rapport du *solde* sur le *volume* de flux au niveau des lieux, définie par D. Thomas (1941), sous l'expression « *d'efficacité démographique* » (*demographic effectiveness*).

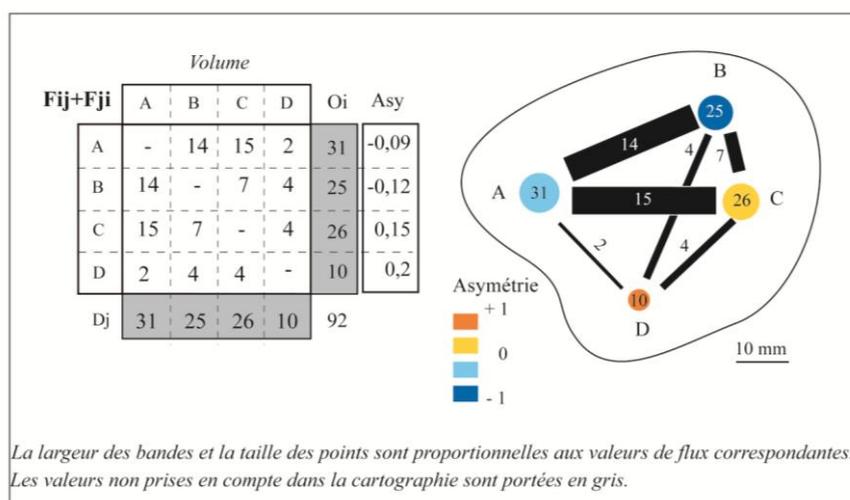
D'après D. Thomas, l'asymétrie ne serait pas tant un indicateur démographique qu'une mesure du caractère unidirectionnel de la migration entre des zones géographiques. Elle serait ainsi particulièrement adaptée à l'analyse des flux agrégés. Cependant, sur le plan cartographique *l'asymétrie* est mobilisée au niveau des lieux et non des liens comme illustré par la partie centrale de la Figure 3-21 (présentée page 136) et cela, indépendamment du fait qu'elle puisse être calculée sur les liens et les lieux, et prendre des valeurs positives et négatives. De plus, la particularité du calcul de l'asymétrie au niveau des lieux mobilise les matrices de volume et de solde, elle ne correspond pas à une somme marginale en ligne : pour le lieu A par exemple, l'asymétrie est de $-3/31 = -0,96$.

Figure 3-21. Illustration de l'asymétrie d'une matrice de flux (Fij)



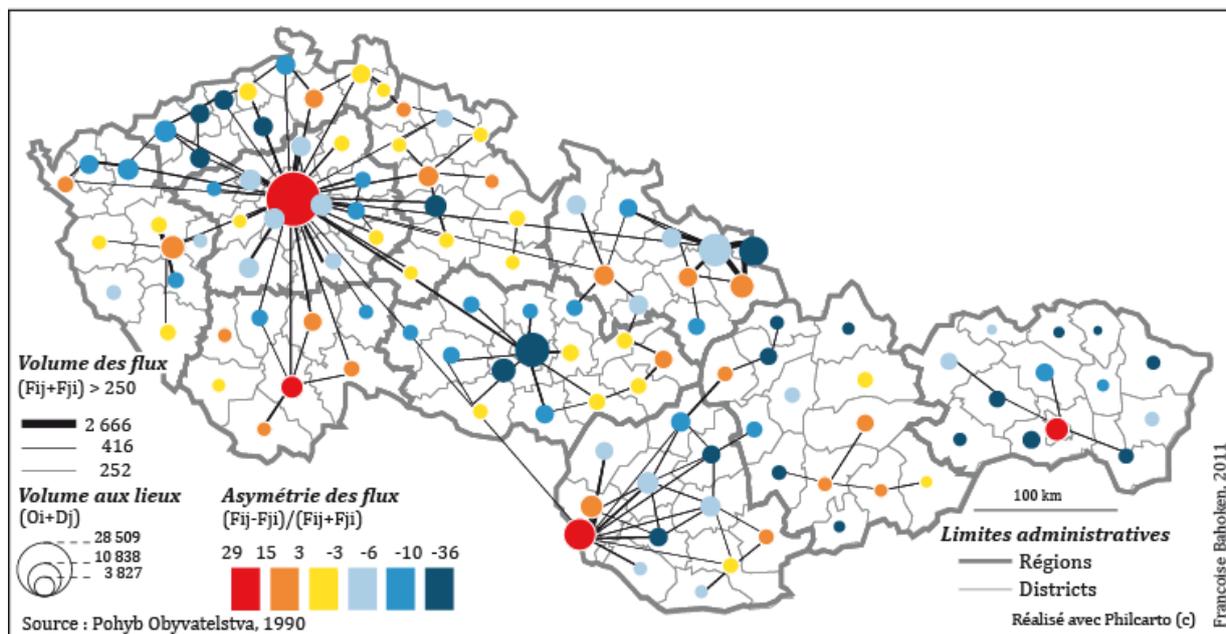
On appréhende souvent l'asymétrie comme une variante de la carte des volumes, qui intègre la représentation des transferts nets au niveau des lieux. L'asymétrie est ainsi représentée au niveau des lieux, par des points proportionnels au volume de flux qui sont assortis d'une coloration discrétisée ; et au niveau des liens, par une bande proportionnelle à la valeur du volume de flux. La formation de cette matrice est présentée sur la Figure 3-21; un exemple de sa cartographie appliquée au flux de l'ex-Tchécoslovaquie est proposé sur la Figure 3-22

Figure 3-22. Illustration de la cartographie classique de l'asymétrie des flux



Ce choix d'appréhender la cartographie de *l'asymétrie* des flux comme une variante de la carte du *volume bilatéral* est effectivement judicieux. Il met en évidence des relations asymétriques où certains lieux émettent beaucoup plus qu'ils ne reçoivent et inversement, tout en décrivant le *volume de flux aux lieux*. De ce fait, la Figure 3-23 laisse deviner un système migratoire régional et fortement polarisé.

Figure 3-23. Cartographie du volume et de l'asymétrie des flux de l'ex-Tchécoslovaquie (1990)



Remarque : le camaïeu de tons bleus (initialement verts) a été modifié en 2016, afin que la charte graphique de la thèse soit respectée.

La cartographie classique de *l'asymétrie* des flux consiste à enrichir la carte du *volume bilatéral*. Les points proportionnels correspondent aux volumes de flux aux lieux : ils présentent une coloration discrétisée en six classes de valeurs opposées qui traduisent l'asymétrie observée du point de vue des lieux. Les valeurs d'asymétrie positive sont teintées d'une couleur chaude tandis que celles qui sont négatives présentent un ton froid, en application du principe de l'opposition colorée des séries divergentes. Des bandes proportionnelles unies sont tracées entre les couples de lieux pour illustrer le

volume des flux bilatéraux. Cette représentation de l'asymétrie des flux traduit efficacement une forme d'attractivité des lieux, en même temps qu'elle permet de deviner partiellement le *transfert net* puisque l'on observe que certains lieux sont figurés en bleu. Si l'on compare cette carte avec la Figure 3-19 b), on observe qu'elle combine efficacement la distribution des valeurs positives et négatives. Son avantage étant d'intégrer les informations qui concernent (F_{ij}^+) et (F_{ij}^-) , elle montre une forme d'attractivité des lieux en même temps qu'elle décrit leur bilan. Si l'on compare cette carte avec celle de la Figure 3-19 c), par exemple, on observe qu'elle combine efficacement la distribution des valeurs positives et négatives des transferts et des bilans nets. Sur le plan sémiologique, la carte est aussi plus riche que les précédentes car elle mobilise trois composantes : des bandes proportionnelles aux volumes bilatéraux des flux ; des points proportionnels aux volumes des flux aux lieux et une coloration discrétisée pour illustrer l'asymétrie des flux.

Il est, bien entendu, toujours possible de faire varier le mode de représentation de ces trois grandes familles de matrices, en agissant sur les aspects cartographiques ou en combinant plusieurs modes de représentation. Il est également possible de faire évoluer cette cartographie de l'asymétrie des flux. Nous montrerons en particulier dans le chapitre 5 que, grâce aux nouvelles méthodes de visualisation des données statistiques, il est désormais possible d'appréhender l'asymétrie des flux comme une variante de la carte des *transferts nets* et non plus seulement comme celle du *volume bilatéral*.

Conclusion du chapitre

La proposition de W. Tobler, de décomposition de la matrice de flux (F_{ij}) initiale en deux composantes, nous a conduits à identifier trois familles de méthodes de cartographie des flux, utilisant la grille d'analyse *liens-lieux* (Grasland, 2009). Ces trois méthodes consistent à représenter soit **les valeurs des flux observés** (F_{ij}), soit les **volumes de flux bilatéraux** (F_{ij}^+) soit encore les **transferts nets bilatéraux** (F_{ij}^-) et, éventuellement, un indicateur de leur asymétrie ($W_{ij} = [(F_{ij}^-)/(F_{ij}^+)]$). Ces trois familles de méthodes correspondent aussi, sur le plan formel, à trois types de matrices de flux qui sont, respectivement, **asymétriques**, **symétriques** et **antisymétriques**.

L'approche fondée sur la décomposition des matrices de flux présente plusieurs avantages. Sur le plan de la sémiologie cartographique, la représentation *liens-lieux* renforce la perception de la structure des flux car ces derniers sont représentés du point de vue des lieux (par les points) et des flux (par des liens) : sur le plan visuel, étant donné que les symboles sont proportionnels aux valeurs, la surface des points est égale à la *largeur* des liens (entrants et/ou sortants). La présence des points de tailles variables met en outre l'accent sur la hiérarchie des lieux et leur importance dans le système d'échanges, sur leur place dans la structure globale des flux, au regard du type de représentation du flux dont il est question (**volume bilatéral**, **transfert net bilatéral**, **asymétrie**). Sur le plan méthodologique, cette approche permet de représenter l'ensemble des parties de la matrice tandis que, compte tenu de *l'effet-spaghetti*, on s'attendait peut-être à ne voir cartographiées que les liaisons entre les lieux, soit une information partielle de la matrice. Enfin, ces résultats nous conduisent à la conclusion selon laquelle notre parti pris cartographique, fondé sur l'approche *liens-lieux* et sur la *décomposition* de la matrice, contribue à améliorer la connaissance thématique des flux concernés : la hiérarchie des lieux qui en résulte pouvant toujours être interprétée pour des flux commerciaux, financiers ou migratoires.

Le Tableau 3-3 récapitule ces trois familles de méthodes ainsi que leurs caractéristiques, selon qu'elles sont appréhendées du point de vue de la matrice (approche matricielle), du point de vue des figurés cartographiques mobilisés¹⁰¹ ou en fonction du processus représenté (appréhendé en termes de compétition ou de coopération, par exemple), au regard de la composante des flux représentée (des flux ou échanges observés, un volume bilatéral ou un transfert net).

Tableau 3-3. Récapitulatif des principales méthodes de cartographie des matrices de flux

Approche matricielle			Approche cartographique (figurés mobilisés)		Approche thématique		Approche théorique (processus représentés)
Type de matrice	Liens (F_{ij})	Lieux (i,j)	lignes	points	Liens	lieux	
Asymétrique	F_{ij}	-	Flèche	Points	Flux	-	
Symétrique	(F_{ij}^+)	(O_i+D_j)	Bande	Points	Volume bilatéral	Volume aux lieux	Coopération
Antisymétrique	(F_{ij}^-)	(D_j-O_i)	Flèche	Points	Transfert net	Solde	Compétition
Complexe	(F_{ij}^+)	(O_i-D_j) (O_i+D_j) et (O_i+D_j)	Bande (ou fleche)	Points discretisés	Volume bilatéral	Asymétrie- Volume	

Le Tableau 3-3 présente les méthodes applicables à n'importe lequel des quatre types de matrices identifiées, en faisant appel à des données internes à la matrice : celles qui portent sur les flux et sur les lieux approximées par les marges de la matrice. De ce fait, il convient de garder à l'esprit que des données exogènes peuvent être requises, en fonction de la thématique analysée et de l'importance accordée à l'espace.

¹⁰¹ Sachant que la *Taille* des figurés est toujours proportionnelle à la valeur du flux dont il est question.

Les marges (O_i , D_j) de la matrice peuvent ainsi être remplacées par des indicateurs descriptifs des lieux ou caractéristiques de leur taille ; les valeurs des flux pouvant être pondérées. Pour les flux migratoires, par exemple, les populations des lieux d'origine et/ou de destination (P_i , P_j) peuvent ainsi remplacer les sommes marginales. Cependant, l'attention est attirée sur le choix de données exogènes qui mérite d'être réalisé avec minutie, par un spécialiste de la question traitée et en fonction de l'objectif de la carte : celles des lieux d'origine pouvant, par ailleurs, différer de celles des lieux de destination. Dans le cas des flux migratoires de type navettes, le terme (P_i) correspond à l'effectif de population résidente, généralement celle des individus de 15 ans et plus ayant un emploi, tandis que le terme (P_j) correspond plutôt à l'offre d'emploi des zones de destination. Il convient aussi de garder en mémoire que l'utilisation de marges ou, plus généralement, de critères liés à la taille des lieux tels que (P_i , P_j) est sensible aux effets de représentativité spatiale et au découpage géographique utilisé lors de leur mesure (Courgeau, 1980 : 135-136).

D'après le Tableau 3-3, l'intérêt des différentes matrices est de porter le regard sur l'une ou l'autre partie du système de flux, ce qui permet d'enrichir l'analyse thématique des flux. La décomposition présente, en effet, l'intérêt de fournir une interprétation thématique des flux. Par exemple, une matrice de (F_{ij}) traduira la dynamique du système, sa cartographie illustrera alors les phénomènes de compétition entre les lieux à l'inverse d'une matrice de type (F_{ji}^+) qui traduira la coopération. Si l'approche est efficace sur les plans théoriques et analytiques, elle ne permet pas pour autant de résoudre le nœud du premier problème de la cartographie de flux qui est celui de **complexité graphique**. L'ensemble des cartes issues de la matrice de flux de l'ex-Tchécoslovaquie, que nous avons réalisées pour illustrer l'approche classique de cartographie de flux, a fait l'objet d'une sélection préalable des valeurs.

Avant d'aller plus avant dans la recherche de solutions permettant de délier ce nœud, nous proposons une rétrospective des méthodes de cartographie des flux : elle fait l'objet du quatrième chapitre.

4^{ième} CHAPITRE. L'évolution des méthodes de cartographie de flux : état de l'art des procédés et des principes de sémiologie graphique

« Les cartes de flux ressemblent à celle des réseaux, mais s'en écartent souvent en raison de la complexité des relations à cartographier. [...] montrer l'ampleur et l'inégalité des mouvements de biens, de personnes, ou d'informations entre un certain nombre de points [...] Au-delà d'un certain nombre de couples, la carte serait illisible, et l'on doit avoir recours à des artifices ».

Roger Brunet (1987).

Introduction du chapitre

La cartographie d'une matrice de flux exprime, par une retranscription cartographique, une manière de voir le monde qui conduit à communiquer une information sur la nature et la structure spatialisées des relations entre des lieux, que celles-ci soient évaluées ou non. Nous avons présenté dans le chapitre 2 trois grands modèles de représentation des matrices de flux. Tous permettent de communiquer une information sur la structure des échanges, qu'elle soit observée depuis les lieux (d'origine et/ou de destination) ou bien des liens, en l'occurrence du point de vue des flux. Nous avons ainsi pu identifier la particularité de la carte de flux, notamment au regard du graphe non planaire, qui réside dans la prise en compte de l'espace, laquelle n'est pas sans effets sur le résultat cartographique. La contrainte forte qui en découle renforce le problème général de la représentation graphique des matrices de flux, que nous avons qualifié d'*effet spaghetti* (chapitre 2). Lorsque la matrice est dense, ce problème de complexité graphique ne semble pas résolu par la seule application d'une méthode de décomposition : les exemples de cartes qui en sont issues (chapitre 3) n'ont porté que sur une sélection de l'information disponible dans la matrice.

Avant d'aller plus loin dans l'énoncé des causes de cette complexité graphique et des solutions qui pourraient être mises en œuvre, objet de la seconde partie, il nous semble intéressant et important de remonter le temps. Cette rétrospective présente un double intérêt : d'une part, celui de suivre l'émergence et l'évolution des méthodes de cartographie de flux, certaines d'entre elles ayant été valorisées quand d'autres sont tombées dans l'oubli ; d'autre part, celui d'analyser la manière dont une sélection d'auteurs a intégré les différentes contraintes liées à la représentation cartographique des matrices de flux. Certains en ont d'ailleurs proposé des comparaisons, en focalisant leur analyse sur *l'objet Carte de flux* (Bunge, 1962 ; Bertin, 1967 ; Dorling, 1991 ; Dent, 1999 ; Zanin, Trémélo, 2003), sans jamais – à notre connaissance – préciser le type de matrice dont il est question. Et pourtant, nous avons vu dans le chapitre 3 que la figure est susceptible de varier en fonction du type de matrice dont il est question.

Si la production cartographique de flux est indéniablement le reflet d'une époque centrée sur la première mondialisation (fin du XIX^{ième} siècle), elle peut être appréhendée comme un produit, c'est

l'objet *Carte de flux*, mais également comme une méthode, une manière de procéder, c'est la cartographie de flux. Le fait de retranscrire sur un fond de carte un processus, la structure d'un système d'échanges, consiste à le figurer par un procédé graphique. En sémiologie graphique, le symbole fait partie de ces procédés graphiques utilisés pour décrire graphiquement un objet, la flèche étant l'un des symboles privilégiés pour représenter le flux.

Le symbole ou figuré est ce qui représente, le *signifié*, et la représentation (la carte) en est un fait, le *signifiant*. Il y a donc plusieurs manières de représenter les flux à l'aide de plusieurs symboles. Le peintre R. Magritte évoque, avec l'expression « *la trahison des images* », la différence conceptuelle entre l'objet représenté – dans notre cas, le flux – et sa représentation par le graphisme ou l'image, par exemple sous la forme d'une flèche. Pour le peintre, la peinture à huile sur toile et pour le cartographe, le crayon (ou plutôt la souris ! sauf pour Philippe Rekacewicz, un célèbre cartographe contemporain). Cette différence entre l'objet et sa représentation élimine toute confusion entre ce qu'on appelle « l'objet représentant » (c'est-à-dire le tracé) ; l'objet lui-même (le flux réel) ; l'acte de représenter (le faire). L'auteur a ainsi analysé deux processus distincts : l'écart entre l'objet, son identification et sa représentation qu'il propose dans des tableaux aux intitulés évocateurs : « *Ceci n'est pas une pipe* » ; « *ceci n'est pas une pomme* » ou encore « *le mot 'chien' ne mord pas* ».

De ce point de vue, la carte de la Figure I-1 (page 13) est-elle une carte de flux ? Qu'est-ce que représenter un flux sur une carte ?

A la différence des objets représentés par R. Magritte, l'objet Carte de flux n'existe pas en tant que tel. On pourrait d'ailleurs reprendre l'aphorisme du comte A. A. S. Korzybski¹⁰² « *Une carte n'est pas le territoire* ». Il s'agit de la représentation conventionnelle d'une construction intellectuelle retranscrite sur un support. La carte est le produit d'une abstraction de la réalité des échanges perçus et représentés par son auteur. D'après B. Debarbieux, « *Elle est représentation d'une réalité inaccessible aux sens, du moins dans l'exercice usuel des compétences naturelles de l'être humain. Elle est le produit de la reconstitution par la pensée, logique et imaginaire, de son référent* » (Debarbieux, 2004 : 608).

Le cas de la cartographie de flux revêt, en outre, un aspect particulier lié à la prise en compte de l'espace et du temps. « *Quand les deux dimensions du plan sont mobilisées pour représenter l'espace, il n'y a plus de dimension plane pour représenter la composante « temps », ce qui pose le problème de la représentation du mouvement en cartographie* » (Bertin, 1967). Cette assertion de J. Bertin explique à elle seule d'une part, le recours tardif au graphique pour illustrer les flux spatialisés et d'autre part, la limitation du nombre d'ouvrages de référence couvrant l'essentiel du sujet que nous avons déjà mentionnée, en comparaison avec l'analyse des distributions de population dans l'espace qui fait l'objet de plusieurs ouvrages et manuels.

Ce quatrième chapitre tente de rassembler une sélection des contributions majeures à la cartographie de flux, qu'elles soient liées au traitement ou à la représentation des données. Il est composé de deux sections. La première est une mise en perspective historique et chronologique des productions cartographiques d'auteurs qui ont fait date, en soulignant leurs apports pour la cartographie de flux et la seconde section est consacrée à la conception de la carte de flux.

La présentation chronologique fait état de six périodes. La première période est focalisée sur les travaux de H.D. Harness qui marquent, de notre point de vue, l'origine de la cartographie statistique de flux. Elle est suivie par l'âge d'or représenté par les travaux des ingénieurs européens de la fin du

¹⁰² Alfred Abdank Skarbeck Korzybski est ingénieur, spécialiste du renseignement et passionné des Sciences de l'homme. Il a beaucoup influencé les milieux artistiques, littéraires et scientifiques, pour avoir fondé ce qu'il nomme « la sémantique générale », définie dans Wikipédia comme une manière de raisonner fondée sur des hypothèses neurologiques, mettent en évidence le fait que l'Homme construit des systèmes de relations, des cartes de cartes, d'ailleurs souvent qualifiées de cartes de flux.

http://fr.wikipedia.org/wiki/S%C3%A9mantique_g%C3%A9n%C3%A9rale [lien vérifié le 07/09/2015].

XIX^{ème} siècle, puis par le tournant pris par la fameuse publication de E.G. Ravenstein sur les Lois de la migration. La quatrième période correspond à l'envol global, expression que nous consacrons à l'émergence d'une cartographie des réseaux et des flux réalisée à l'échelle mondiale, à la faveur d'un courant général qualifié par E. de Martonne « *d'ère aérienne de la géographie* ». La cinquième période est celle de l'apogée de la cartographie de flux : elle est caractérisée par le transfert effectif vers la géographie, a fortiori la cartographie statistique, de concepts mis en œuvre dans d'autres disciplines, en particulier en physique. Cette période est souvent considérée comme celle du « *tournant spatial* » ou du « *tournant quantitatif* » de la géographie humaine. La sixième et dernière période est plus contemporaine : elle s'inscrit dans l'ère de l'informatisation, du *web 2.0* qui a conduit à la généralisation du support écran et des techniques graphiques associées. Elle est à l'origine de d'acquis d'ordres graphiques et esthétiques obtenus dans le champ de la visualisation interactive des réseaux susceptibles de renouveler la figure des flux.

Cette rétrospective est présentée de manière chronologique dans l'objectif de mettre en lumière l'évolution de la pratique cartographique ayant contribué à l'inclure à l'interface de deux champs distincts, mais complémentaires : celui du traitement mathématique des données et celui de la conception graphique. Ces deux champs correspondent aux deux sections principales de ce chapitre. Dans la première section, la présentation de cartes anciennes intègre nécessairement les questions liées au traitement des données, c'est-à-dire à la génération des valeurs de la matrice dont est issue la carte. Celles-ci seront mises en perspective avec la typologie en trois grandes familles de méthodes de cartographie de flux proposée dans le chapitre 3¹⁰³. Cette typologie se voulant générale et généralisable, nous souhaitons la valider en l'appliquant à des cartes de flux anciennes. Pour ce faire, un court texte suggérant le type de matrice utilisé pour la représentation concernée sera proposé au pied des figures. Dans la seconde section, l'attention est focalisée sur les différents procédés et principes de sémiologie cartographique de flux, mis en œuvre lors de l'étape de conception de la carte. Cette section présente également les contraintes graphiques et les variantes qui ont pu être proposées.

4.1 Rétrospective (1786 -) des méthodes de cartographie de flux

L'idée qui consiste à représenter graphiquement les données quantitatives trouve son origine au XVIII^{ème} siècle, notamment dans les travaux de l'économiste W. H. Playfair (Tufté, 1983), considéré comme l'inventeur de la statistique graphique, pour avoir proposé la Figure 4-1.

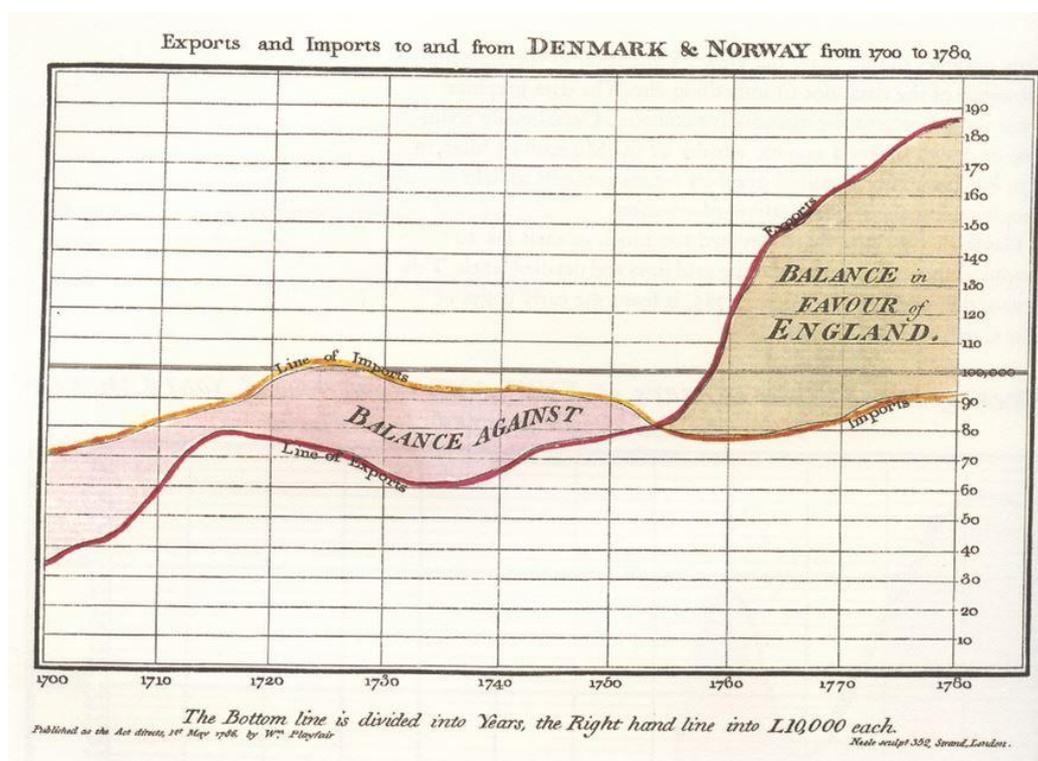
Pour la France, d'après G. Palsky, c'est C. de Fourcroy (1782) qui est considéré comme précurseur dans l'utilisation graphique de pictogrammes proportionnels (Palsky, 2004b : 587-588). Quoiqu'il en soit, la nouveauté du graphique de W. Playfair, dont on notera qu'il porte sur des flux commerciaux, réside dans la double échelle proposée¹⁰⁴, en abscisse. A la suite de W. Playfair, il faudra ensuite attendre la seconde moitié du XIX^{ème} siècle pour pouvoir disposer des éléments conceptuels et méthodologiques, encore d'actualité aujourd'hui, pour représenter des flux sur une carte.

En effet, les premières cartes thématiques apparues, d'après G. Palsky (1996, 2004b), aux XVII^{ème} et XVIII^{ème} siècle ont, en effet, d'abord porté sur la représentation des stocks, sous la forme de cartodiagrammes superposés à des cartes topographiques. Au XIX^{ème} siècle émerge le champ de la cartographie statistique à des fins d'illustration des phénomènes démographiques et sociaux, champs développé essentiellement par des administrateurs, des ingénieurs et certains médecins, notamment pour illustrer des phénomènes de diffusion d'épidémies.

¹⁰³ Pour généraliser le propos, nous parlerons de cartographie d'une matrice de type (Fij), (Fij⁺) et (Fij).

¹⁰⁴ Une description de ce graphique est accessible à l'adresse suivante : http://fonds-ancien.ensmp.fr/Expositions/Graphiques/Pages/deux_dimensions.html [Lien vérifié le 22 mai 2010].

Figure 4-1. Le premier graphique statistique sur des flux (W. Playfair, 1786)



Source : Fonds ancien, Bibliothèque de Mines ParisTech. Paris.

La matrice dont ce graphique fournit une illustration correspondrait à la famille (Fij). L'auteur aurait utilisé les valeurs (Fij) et leur transposée (Fji) puisque la nature des flux varie en fonction de leur orientation, ils correspondent respectivement aux courbes rouges (pour les exportations) et jaune (importations). La zone de différence entre ces deux courbes de valeurs, portée en bleu sur le graphique, correspondant au solde bilatéral.

G. Palsky (2004c) précise ainsi que la première carte statistique est conçue en 1826 par C. Dupin, membre de l'Institut et de l'Académie des Sciences, selon la méthode dite Dupin. Réalisée en noir et blanc, la *Carte figurative de l'instruction populaire de la France* est une représentation de la scolarisation de la France. Les départements sont symbolisés par un aplat de couleurs dégradées : du noir, pour les départements où la scolarisation est la plus faible au blanc, pour représenter le nombre d'élèves par habitant. A l'époque, les méthodes introduites sont d'abord celles des cartes par points (1830), puis des cartes en aplats ou choroplèthes, des cartes en isolignes ou isoplèthes (1843), enfin des cartes de flux (1837, 1845).

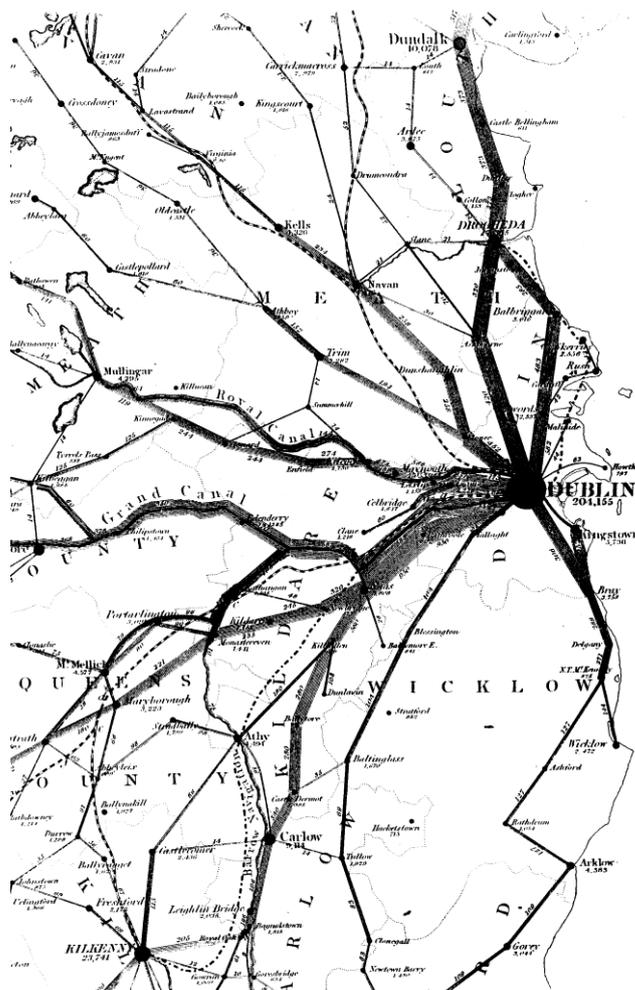
Les méthodes de cartographie de flux sont développées dès 1837, simultanément par des ingénieurs de plusieurs pays européens (Palsky, 2004a, 2004b). « Dès 1830, on passe à la cartographie des ingénieurs car eux-mêmes avaient besoin de connaître la répartition de la population pour décider par exemple des futurs tracés des voies ferrées » (Désiré, 1991 : 445, cité par G. Palsky), ces derniers nécessitant l'intégration de données variées liées aux contextes physiques, techniques et géographiques environnant l'infrastructure (Palsky, 1996, 2004b).

On compte parmi ces ingénieurs H. D. Harness pour la Grande-Bretagne dont la première carte porte sur des flux par voie routière est publiée en 1837, et C.-J. Minard (1845-1869), ingénieur français des Ponts et Chaussées qui a entamé ses travaux de cartographie statistique par une carte de flux, également par voie routière. Ces deux auteurs sont des figures emblématiques de la représentation des échanges, à une époque de fort développement de la cartographie statistique. D'après A. Robinson (1955, 1982), c'est entre 1835 et 1855 que les techniques de représentations cartographiques de population, encore en vigueur aujourd'hui, « [...] *des effectifs de populations, des distributions, des densités et des mouvements* » (Robinson, 1955 : 1), ont été introduites. L'origine de la cartographie spécifique des flux remonterait ainsi à la publication en 1844, de la carte réalisée quelques années plus tôt par H. D. Harness (1836).

4.1.1 L'origine (H. D. Harness, 1836)

La première figure marquante de la cartographie de flux fait partie d'un ensemble de six cartes réalisées en 1836, par H.D. Harness, pour accompagner le second rapport de *l'Irish Railway Commissioners* publié en 1838. Cette carte s'appuie sur la proposition graphique de W. Playfair d'utiliser des symboles proportionnels pour représenter des quantités absolues. C'est ainsi que l'auteur réalise la première carte statistique mobilisant la *largeur* de figurés linéaires (les réseaux de routes), leur *longueur* étant celle de l'infrastructure routière.

Figure 4-2. La première carte de flux (H.D. Harness, 1836)



Source : A. Robinson, 1955.

La matrice dont il est question serait de type (Fig⁺). Les données étant issues de comptages routiers effectués par tronçons, l'auteur représente le volume bilatéral des flux (en l'occurrence le trafic routier) de chacun des tronçons, sans distinguer leur orientation, en faisant varier leur *largeur*.

La carte de la Figure 4-2 représente, en effet, les flux de passagers – un trafic routier – en Irlande mesurés aux différents tronçons routiers au départ de Dublin. Elle met en évidence la distribution relative du nombre de passagers sur les voies routières, dans les différentes directions, ainsi que la décroissance des effectifs de passagers à mesure que l'on s'éloigne de Dublin. Il est important de noter qu'il s'agit de la première représentation cartographique qui fasse état du caractère proportionnel des du linéaire du réseau au trafic et ce, avant même que ne ce soit appliquée cette méthode dans les cartes d'états. La carte introduit par ailleurs le procédé *des flux affectés sur réseau* qui consiste à plaquer la variable visuelle *Taille* directement sur l'infrastructure physique qui supporte la représentation. C'est ainsi que la *largeur* du tronçon est rendu proportionnelle au volume de passagers, leur *longueur* ne pouvant être modifiée car le réseau est planaire. Nous reviendrons sur ces aspects liés au tracé des dimensions du figuré qui décrit les flux, dans les sous-sections 4.2.2 (page 187) et 4.2.3 (page 189) ; aussi dans le chapitre 6. Ce procédé est toujours en vigueur pour représenter des flux matériels, qui s'expriment sur un réseau tels que le trafic routier.

La représentation cartographie de ces flux par voie routière sur un fond de carte est une nouveauté. La pratique de l'époque consistait plutôt à réaliser des diagrammes pour illustrer le temps et des cartes pour représenter les rapports à l'espace. G. Palsky précise que le tournant est pris après 1800, avec le développement de cartes de flux de toutes sortes qui « [...] *inscrivent le temps dans l'espace : cartes des mouvements des armées, de la diffusion des épidémies, des flux de marchandises ou de voyageurs* » (Palsky, 1996).

Le support statique n'est alors pas considéré comme impropre à la représentation d'un phénomène dynamique, en particulier le mouvement. Au même moment, C.-J. Minard, à la fin d'une première carrière d'ingénieur et d'enseignant, entame un travail important de cartographie statistique qui s'étale de 1844 à 1877 et fera date dans l'histoire de la représentation des flux, des trafics et aussi des mouvements. C'est l'âge d'or de la cartographie de flux, l'âge de l'enthousiasme de la cartographie statistique de G. Palsky (1996).

4.1.2 L'âge d'or (C.-J. Minard, 1844-1877)



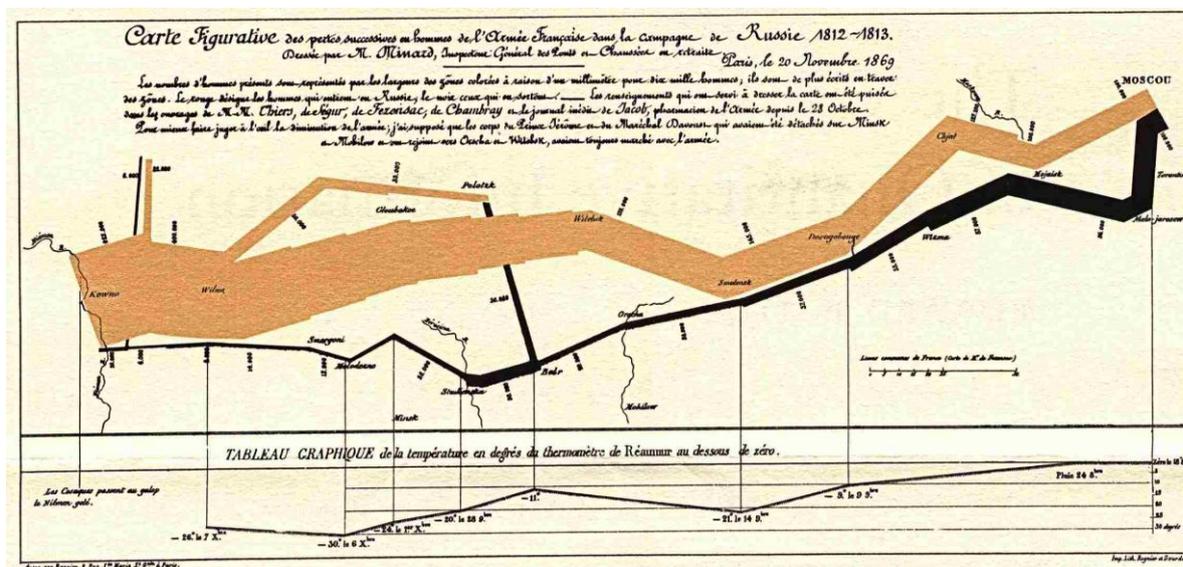
Le plus célèbre des documents cartographiques français sur les flux est sans aucun doute la *Carte figurative des pertes successives en hommes de l'armée française dans la campagne de Russie, 1812-1813* réalisée en 1869 par C.-J. Minard, ingénieur français de l'Ecole des Ponts et Chaussées (voir Figure 4-3) alors retraité. Cette figure met en œuvre le procédé des flux distributifs caractéristique de l'auteur.

D'après E. Tufte, la Figure 4-3 représente, « [...] *Probablement le meilleur graphique statistique jamais réalisé* » (Tufte, 1983). De notre point de vue, si cette figure est importante, c'est parce qu'elle permet de représenter un mouvement et non seulement un flux (nous y reviendrons dans le chapitre 6), pour deux raisons principales. Tout d'abord, parce que le cheminement exact entre les lieux est connu de l'auteur, ce qui n'est pas toujours le cas pour des données agrégées ; ensuite parce que la figure mobilise un procédé graphique inédit, caractéristique de l'auteur qui en fait toute la richesse sémiologique.

Cette image du mouvement des troupes napoléoniennes est aussi la dernière que l'auteur réalisa (Mac Eachren, 1979 : 28), ce qui explique probablement sa richesse, sur le plan méthodologique. Les

apports¹⁰⁵ de C.-J. Minard à la cartographie de flux sont tels que nous en proposons un classement thématique. Nous structurons ainsi notre propos autour d'une sélection de ses cartes figuratives, révélatrices des différentes astuces et procédés mis en œuvre que nous considérons comme constitutifs de l'âge d'or de la cartographie de flux. L'essentiel des travaux de l'auteur pouvant désormais être consulté en ligne¹⁰⁶ sur le site dédié au patrimoine de l'Ecole Nationale des Ponts et Chaussées, Ponts Paris Tech (ENPC).

Figure 4-3. La carte de C.-J. Minard (1869) sur la campagne napoléonienne de Russie



Source : C.-J. Minard, 1869.

La matrice dont est issue cette représentation serait de type (Fij), l'auteur souhaitant marquer la différence entre la valeur des flux aller (qui représente la quantité de soldats), en brun sur la figure, par rapport à celle des flux retours de la campagne, représentés en noir. A noter toutefois que, de notre point de vue, le tableau ne serait pas une matrice de flux stricto sensu, mais un tableau de déplacements qui décrirait le cheminement de cette armée de soldats par différents lieux séparés par des tronçons. D'ailleurs, les lieux par lesquels ces soldats sont passés à l'aller de cette campagne russe ne correspondent pas particulièrement à ceux par lesquels ils sont passés au retour.

L'analyse d'une partie de la production de C.-J. Minard nous a, en effet, conduites à identifier quatre principes majeurs de la représentation des flux encore en vigueur de nos jours. Le premier est celui de la généralisation de la variable visuelle *Taille* pour la représentation des quantités déplacées. Le second porte sur l'utilisation de la variable *Couleur* (ou teinte) pour discriminer les échanges et permettre une lecture à plusieurs échelles de la carte. Le troisième principe, qui tient à l'époque davantage de l'astuce, a trait à la gestion de la superposition des figurés qui illustrent les relations entre les lieux et le quatrième consiste en la généralisation aux flux immatériels, d'un procédé des flux affectés sur réseau, qui prend la forme d'un procédé distributif plus connu sous le nom de diagramme Sankey (nous y reviendrions dans la section 4.1.2.4, page 151).

¹⁰⁵ D'après A. Robinson (1982), la production cartographique de C.-J. Minard a atteint « [...] un niveau de sophistication extraordinaire pour son époque », contribuant à étendre son domaine d'application depuis le champ des transports vers celui des Sciences humaines et sociales.

¹⁰⁶ Une sélection de cartes et graphiques réalisés par C.-J. Minard est disponible sur la bibliothèque numérique de l'ENPC, accessible en ligne à l'adresse suivante : <http://patrimoine.enpc.fr/collections/show/12>

4.1.2.1 L'utilisation de la Taille pour la représentation des quantités déplacées

La particularité des cartes de flux réalisées par C.-J. Minard est qu'elles sont pratiquement toutes fondées sur le procédé des flux affectés sur réseau, mis en œuvre par H.D. Harness (1836) sur des tronçons routiers. L'auteur transfère ce procédé à la cartographie de flux immatériels origine-destination, ce qui constitue une première nouveauté. Il contribue par ailleurs à généraliser la variable visuelle « taille »¹⁰⁷, ce qui constitue une innovation à l'époque. G. Palsky (2004b) considère d'ailleurs l'auteur comme le « Playfair français » en référence aux diagrammes et cartes qu'il produit. La première de ses Cartes figuratives réalisée sur ce modèle dès 1845, est la *Carte de la Circulation des voyageurs par voitures publiques sur les routes de la contrée où sera placée le chemin de fer de Dijon à Mulhouse (Mars 1845)* »¹⁰⁸.

C.-J. Minard commence, comme H.D. Harness, par représenter des flux affectés sur une infrastructure. Il étendra ensuite le procédé aux flux immatériels : les liens entre les lieux sont désormais proportionnels aux quantités, tout comme les stocks aux lieux. Cette idée, plus que commune aujourd'hui, est particulièrement innovante à l'époque. G. Palsky (1998) précise ceci : « *Autour des axes routiers, Minard construit une bande d'épaisseur proportionnelle au trafic : un millimètre pour mille voyageurs transportés en une année* » (Palsky, 1998 : 53), ce qui n'est pas sans rappeler la méthode mise en œuvre pour réaliser la carte de la Figure I-1-1 (page 13).

D'après A. Mac Eachren, « [...] en travaillant de manière indépendante en France, [C.-J. Minard] a produit la première carte qui ait utilisé [la méthode des] des cercles proportionnels pour illustrer une information autre que la [distribution de] population. Funkhouser affirme que Minard a publié la première de ses cartes en 1851, mais que le sujet de la carte n'était pas mentionné. Robinson présente, en exemple, l'extrait d'une carte en symboles proportionnels publiée par Minard [...] les tonnages aux ports les plus importants du monde » (Mac Eachren, 1979 : 24), représentant le volume de flux en différents lieux.

4.1.2.2 L'utilisation de la Couleur (la teinte) pour discriminer le type de flux

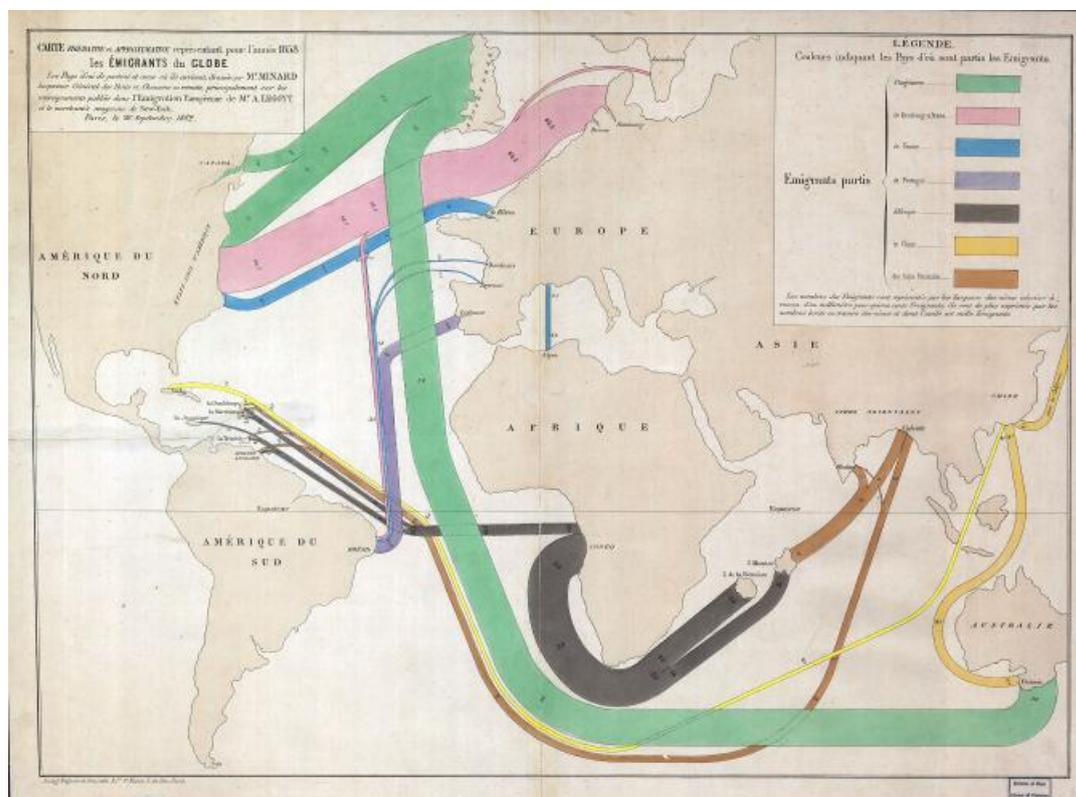
La seconde innovation apportée par les travaux de C.-J. Minard porte sur l'utilisation de la couleur, au sens de teinte, sur des données qualitatives. La teinte permet à l'auteur, dans un premier cas d'utilisation, de discriminer les types de parcours. Ainsi, dans la *Carte de la Circulation des voyageurs par voitures publiques sur les routes de la contrée où sera placé le chemin de fer de Dijon à Mulhouse (Mars 1845)*, C.-J. Minard effectue une distinction subtile : « [...] en gris et noir les parcours partiels du parcours total (entre les points extrêmes du réseau) » (Palsky, 1998 : 53) ; tandis que dans sa *Carte figurative et approximative des tonnages des vins, spiritueux et qui ont circulé en 1857 sur les voies d'eau et de fer de l'Empire français*, réalisée en 1860, les couleurs utilisées distinguent les différentes voies de transport selon qu'elles sont maritimes, terrestres ou d'exportation. Il autorise ainsi une double lecture des flux ainsi représentés : ceux qui sont internes (ou intérieurs) et ceux qui sont externes (ou extérieurs), franchissant la frontière nationale.

Dans un second cas d'utilisation, la couleur permet à C.-J. Minard de discriminer les types de flux en fonction de leur provenance. C'est notamment le cas dans la *Carte figurative et approximative représentant pour l'année 1858 les immigrants du globe* réalisée en 1862 présentée sur la Figure 4-4.

¹⁰⁷ La variable visuelle *Taille*, mise au point par W. Playfair ne sera introduite en France, d'après G. Palsky, qu'en 1844, « [...] par un obscur ingénieur ordinaire, [...] Bollain qui utilise des cercles proportionnels à la population agglomérée » (Palsky, 2004b : 598)

¹⁰⁸ Source : Bibliothèque des PontsParisTech, Campus Descartes, Université Paris-Est, Champs-sur-Marne, 1845.

Figure 4-4. La gestion de la couleur, selon C.-J. Minard (1862)



Source : C.-J. Minard, 1862.

Cette figure serait réalisée à partir d'une matrice symétrique de type (F_{ij}^+) , qui traduit le volume bilatéral des échanges, soit une valeur non orientée. Cependant, étant donnée la thématique de cette image qui décrit des flux d'immigrés, l'auteur a souhaité représenter la provenance des migrants en recourant à une palette de couleurs pour le moins symbolique : les émigrés partant de la France sont en bleu quand ceux qui quittent la Chine sont en jaune...

Cette carte présente enfin l'avantage d'autoriser un double niveau de lecture. En effet, outre la vision d'ensemble proposée, une connaissance plus fine du système migratoire est rendue possible par la mention des valeurs des flux sur les figurés.

Afin de faciliter la lecture d'ensemble, ce ne sont pas les pays qui ont été dotés d'une teinte mais plutôt leurs liens. La correspondance entre couleur des liens et pays émetteurs est précisée en légende. Cette mise en œuvre de la couleur permet une lecture instantanée des flux à la fois du point de vue des lieux (états d'origine et de destination) que de celui des flux, par les liens rendus proportionnels aux effectifs de migrants.

La *Couleur* et la *Taille* ne sont pas les seules variables visuelles mises en œuvre par C.-J. Minard. Aux côtés des variables définies par J. Bertin (1967), l'auteur intègre *i*) la *localisation* en x et y ; *ii*) l'*orientation*¹⁰⁹ qui indique ici la direction ou le sens des flux ; *iii*) la *Taille* et *iv*) la *Couleur* ainsi que

¹⁰⁹ C. Zanin (2013) nous fait souligner que l'orientation n'est pas une variable visuelle attribuée à un objet linéaire, mais plutôt à un poncif (Bertin, 1967).

d'après E. Tufte (1983), le temps (pour le moment des événements) et la température, telle qu'elle est représentée, par exemple, le long des trajectoires (voir par exemple, sur la Figure 4-3147) ce qui porte à six le nombre de variables mobilisées par C.-J. Minard. Outre une bonne maîtrise du langage graphique, l'auteur y ajoute une astuce sur la gestion de la disposition des figurés, qui correspond à la *surcharge graphique de l'effet spaghetti*.

4.1.2.3 La disposition des figurés pour gérer leur superposition

La troisième astuce mise en œuvre par C.-J. Minard sur ses cartes figuratives porte sur la gestion du chevauchement des liens, c'est-à-dire sur leur superposition. Sur la Figure 4-4, les liens les plus importants en valeur sont placés au dessus de ceux qui le sont moins, mais ce procédé n'est pas très courant. En général, pour faciliter la lecture des cartes de flux, il est d'usage de placer les liens les plus fins au dessus des plus gros (Beguin, Pumain, 1994 ; Zanin, Trémélo, 2003). Le choix de l'auteur est, bien entendu, parfaitement justifié puisqu'il s'agit, pour lui, de donner à voir et au premier plan, les échanges les plus importants. C'est d'ailleurs la solution préconisée par W. Tobler (1987) que nous développons dans la section de ce chapitre consacrée à la disposition des flèches (voir 4.2.2.4, page 189).

La disposition des plus gros flux au premier plan est renforcée par leur aspect translucide qui donne une visibilité aux petits flux qui sont disposés sur les plans inférieurs ; cet aspect permet aussi de conserver la trace des petits flux qui est matérialisée par des tirets placés à l'endroit même du chevauchement entre petits et gros flux (voir Figure 4-4). Il convient de noter également¹¹⁰ l'arrangement¹¹¹ visuel réalisé C.-J. Minard, qui pourrait par ailleurs faire passer la carte du statut de carte à celui de graphique. Pour garantir la visibilité d'ensemble des flux, les contours des continents ont été modifiés par endroits et généralisés à l'extrême : les superficies ne sont donc plus correctes, ce qui pourrait lui être reproché.

La gestion de la disposition des flux selon leur taille, telle qu'elle est réalisée par C.-J. Minard, est caractéristique d'un enchevêtrement de conduits d'énergie. Ainsi, le tracé des routes maritimes qui apparaît dans cette Figure 4-4 résulte très certainement d'un choix lié à des contraintes d'ordre graphique plus qu'une localisation exacte des routes car, théoriquement, il n'y a pas un seul chemin pour aller d'un point à un autre dans le contexte d'une cartographie de flux origine-destination. Cela étant, le procédé de représentation des flux traduit une certaine dynamique, par la courbure des liens, en même temps qu'il représente de façon très efficace les quantités émises et reçues par les lieux.

Dans l'ensemble, les cartes figuratives de C.-J. Minard apparaissent comme une synthèse graphique qui autorise différents niveaux de lecture, en raison de l'additivité des valeurs des flux sortants ou entrants. Sur la Figure 4-4, par exemple, le total des flux émis en provenance de la région du Congo, en Afrique centrale, prend deux directions principales : le sud du continent vers les Indes qui concentre l'effectif de migrants le plus important et l'Amérique latine aux effectifs moindres. Ce dernier flux se subdivise ensuite à l'approche du continent américain en deux flux secondaires : l'un vers la Guyane Anglaise, la Trinité et la Jamaïque et l'autre vers les Antilles françaises, la Guadeloupe et la Martinique.

Ce procédé particulier de représentation en bandes proportionnelles qui se scindent en fonction de la direction de ses ramifications a vu son aboutissement conceptuel et sa généralisation dans les travaux du capitaine M. H. P. R. Sankey, ingénieur irlandais et contemporain de C.-J. Minard.

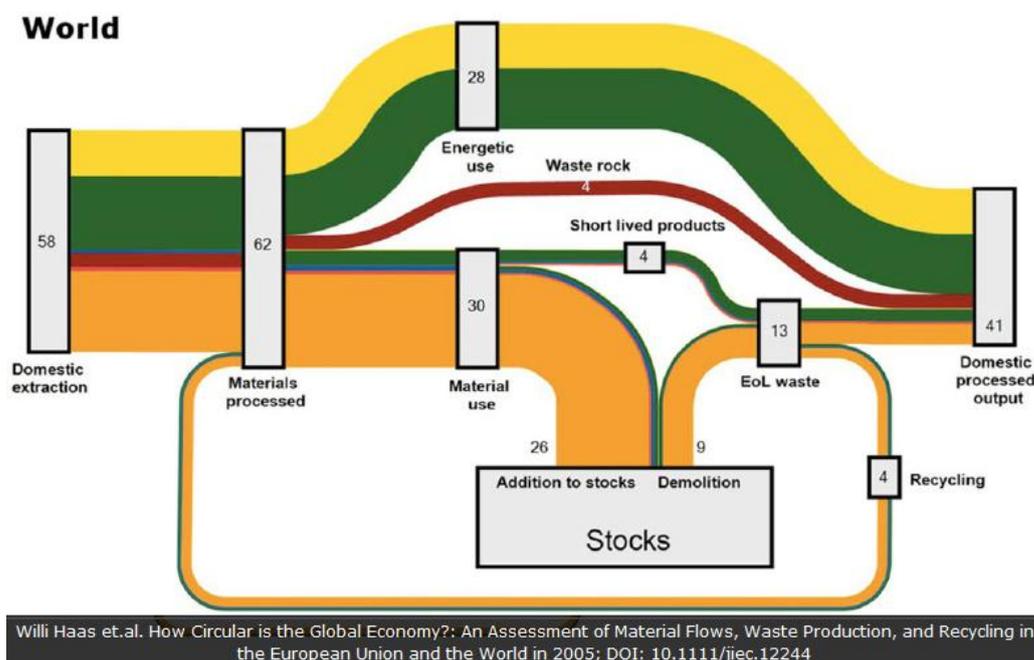
¹¹⁰ Une analyse des cartes de C.-J. Minard est disponible dans les archives du blog *Cartographia : mapping the world around you*, à l'adresse suivante <http://cartographia.wordpress.com/category/charles-joseph-minard/> (vérifié le 21/07/2015).

¹¹¹ Des observateurs minutieux, inconnus, ont pointé les petits arrangements de C.-J. Minard avec les frontières des continents.

4.1.2.4 La généralisation du procédé de Minard (M. Sankey, 1898)

M. Sankey doit son immense succès à la publication d'un graphique particulier dans une minute de l'organisme des Ingénieurs en mécanique en 1898. Ce diagramme compare l'efficacité énergétique d'un moteur à vapeur classique et celle d'un moteur à vapeur sans perte d'énergie, à l'aide de bandes orientées d'intensité proportionnelle et de couleurs variables selon la matière transportée. Le succès du graphique est tel que, dans de nombreuses disciplines, il est désormais d'usage de parler de diagramme-Sankey alors que le principe a pourtant été développé à la suite des travaux de C.-J. Minard (!). Il est de conception simple même si sa mise en œuvre manuelle est moins aisée qu'il n'y paraît : les flux entrants et sortants sont représentés par des bandes d'intensité proportionnelle, parfois orientées par des têtes de flèches.

Figure 4-5. Illustration d'un diagramme Sankey (2005)



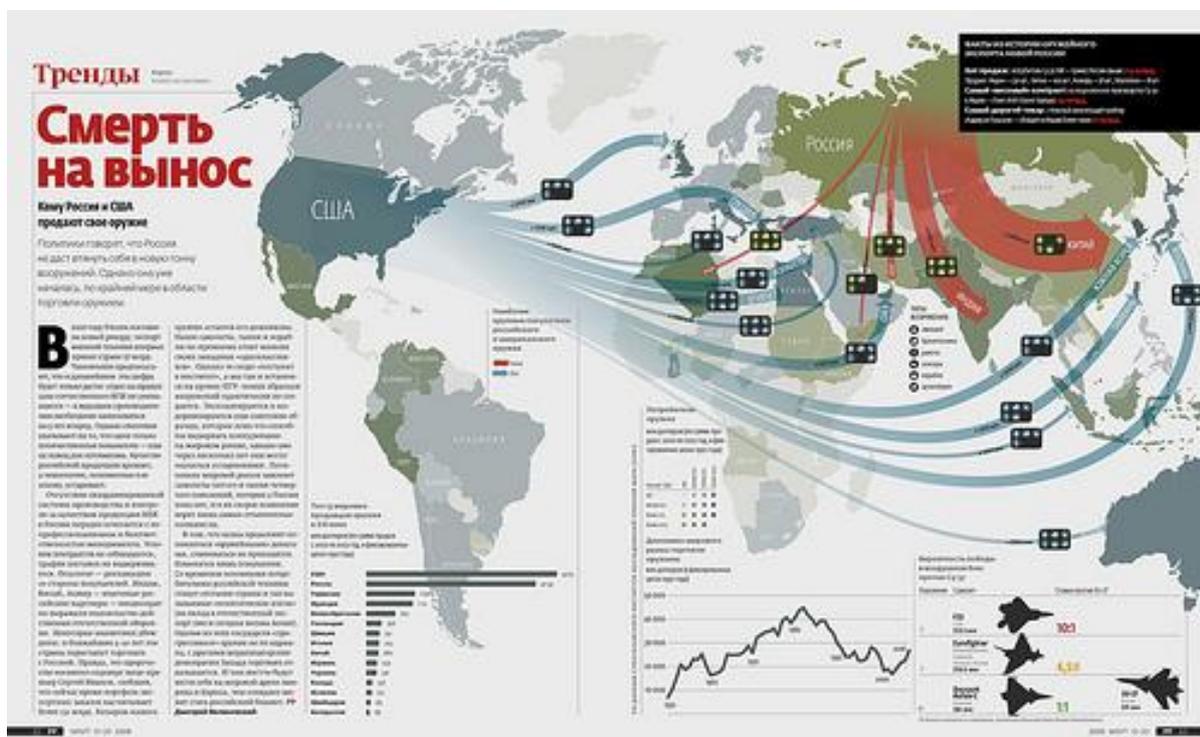
Source : <http://www.sankey-diagrams.com/> [Lien vérifié le 21/07/2015]

Les diagrammes *Sankey*, généralement élaborés à partir d'une matrice de type (Fij), sont souvent assortis d'une couleur afin de représenter le transfert d'une matière ou sa direction. La raison tient au fait que le procédé ne privilégie pas la réalisation de pointes pour signifier l'orientation. En effet, au droit des zones d'origine et/ou de destination, les bandes sont totalement agrégées visuellement et non plus simplement associés ou accolées, à la manière de C.-J. Minard. Si les diagrammes *Sankey* ont surtout une application dans les Sciences Pour l'Ingénieur (SPI), pour illustrer les réseaux physiques, ils ont parfois fait l'objet d'applications cartographiques (voir Figure 4-6).

Il est désormais d'usage courant de ne représenter que la valeur (Fij) des flux ou leur seule composante positive (Fij⁺), qui correspond au volume bilatéral, par des bandes proportionnelles aux quantités échangées ; d'égale épaisseur lorsque seule la visualisation de la quantité importe (et non l'orientation). Dans ce cas, il s'agit davantage de la représentation d'un réseau de relations à partir d'une carte qui n'autorise qu'une comparaison des flux échangés. Plus d'un siècle plus tard, les

travaux de C.-J. Minard sont toujours mentionnés et largement cités. Ils font d'ailleurs l'objet de nombreux développements¹¹².

Figure 4-6. Exemple de carte réalisée selon le principe du diagramme Sankey



Source : Russian Reporter Magazine¹¹³, septembre 2008 (<http://www.novichkov.net>)

La méthode des flux affectés sur réseau, pour la cartographie de flux immatériels, a en effet, récemment fait l'objet d'une automatisation. L'application intitulée *Flow Map Generator*®, développée en 2013 par B. Simantel pour la version 10x du logiciel ArcGis®, permet désormais de réaliser une carte de flux en s'inspirant du procédé de C.-J. Minard. D'après l'auteur, l'outil génère « [...] une carte de flux distributive depuis un lieu d'origine vers un voire plusieurs lieux de destination »¹¹⁴.

Entre temps s'impose de façon quasi universelle le procédé graphique de la flèche pour représenter des relations entre des lieux. Celui-ci est apparu, pour la première fois, sur une carte de flux en 1885. Son avènement marque, de notre point de vue, un tournant dans la représentation de flux tant au niveau du procédé (la flèche) qu'à celui du processus représenté : un solde bilatéral (Fig.).

¹¹² Voir notamment les démonstrations proposées sur la page personnelle de M.-J. Kraak : <http://www.itc.nl/personal/kraak/>

¹¹³ <http://www.flickr.com/photos/27398292@N06/2995018860/> [Lien vérifié le 10/03/2010].

¹¹⁴ Le code source de l'application est disponible en libre accès sur le site de ArgGis : <http://www.arcgis.com/home/item.html?id=3bb350838ee044979d1f583b116cb8c1> ; il est accompagné d'un jeu de données tests : <http://www.arcgis.com/home/item.html?id=9d48ce9b6837492e982e980402d3ed11>. [Liens vérifiés le 02/08/2015].

4.1.3 Le tournant (E.G. Ravenstein, 1885)

La plus ancienne carte civile de flux qui figure des flèches semble être la carte intitulée *Currents of migrations*, qui porte le numéro 5 de l'article *Migrations Laws (Les Lois sur la migration)* de E.G. Ravenstein (1885, 1889). Cette carte s'inscrit dans la démarche particulière de E.G. Ravenstein, géographe et cartographe allemand de la *Royal Geographical Society* de Londres. L'auteur adopte une démarche intuitive, en réponse à W. Farr (1876), à l'époque Président de la *Royal Statistical Society* de Londres (1871-1873), qui affirme que les migrations de populations humaines se produisent de manière non structurée, sans loi apparente (Ravenstein, 1885 : 167).

La réponse porte alors sur la formulation de sept lois¹¹⁵ sur les migrations humaines (voir Encart 4-1) que E.G. Ravenstein illustre par une série de cartes statistiques.

Encart 4-1. Les lois de la migration (Ravenstein, 1885)

- 1) *La plupart des migrants n'effectuent qu'une migration de courte distance. Ceux qui effectuent une migration de longue distance vont de préférence dans des centres de commerce ou industriels.*
- 2) *Il résulte alors un déplacement progressif de la population en direction des centres.*
- 3) *Chaque courant migratoire produit un contre courant compensatoire.*
- 4) *Les personnes du milieu urbain migrent moins que celles du milieu rural.*
- 5) *Les femmes sont mieux représentées dans les flux de courte durée.*
- 6) *La migration augmente avec le développement de la technologie.*
- 7) *Le principal motif des flux migratoires est le désir de l'homme d'améliorer son statut sur le plan matériel.*

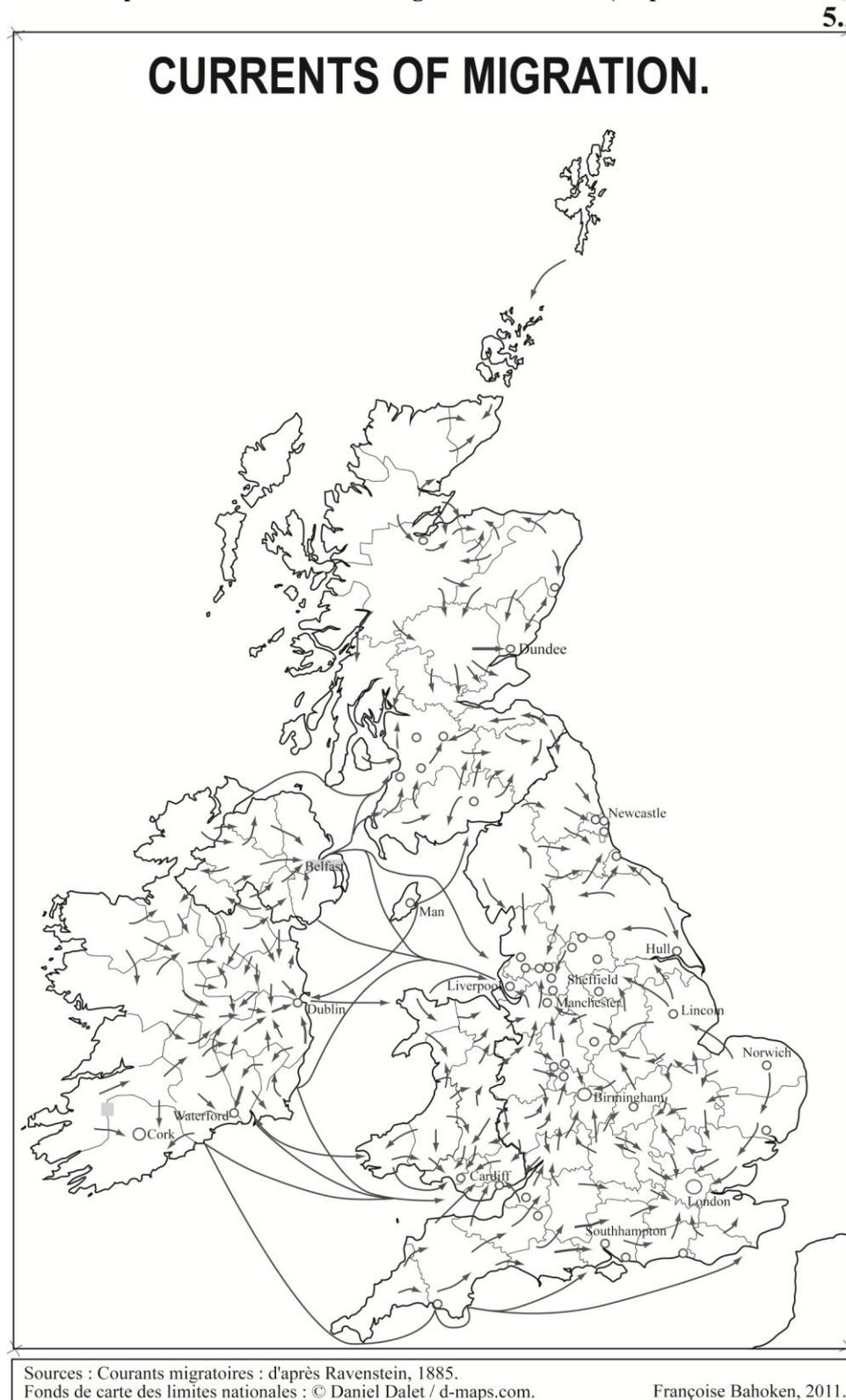
Source : Ravenstein (1885).

Douze cartes des mobilités britanniques, réalisées à l'échelle des comtés, illustrent les lois de la migration, parmi lesquelles la carte n°5. Difficile à interpréter au premier abord, cette carte présente pourtant un intérêt certain pour notre recherche (Bahoken, 2013), elle est présentée sur la Figure 4-7.

Il est important de noter que cette carte n°5 est la seule des douze documents réalisés par l'auteur, qui représente des flux migratoires à l'aide de flèches. Elle a été publiée en l'état, sans aucun élément classique d'habillage (légende, échelle). Elle s'appuie sur un fond de carte extrêmement précis du contour des îles britanniques : on y distingue un maillage administratif interne supposé être des limites de comtés. Certaines localités sont représentées par des cercles blancs, parfois assortis d'un toponyme. Un ensemble de flèches, dont on ne sait pas exactement à quoi elles correspondent, se superpose au maillage. Elles peuvent toutefois être organisées visuellement en deux groupes. Le premier groupe de flèches correspond à celles qui sont courtes, fines et parfois fortes : elles représentent, selon les cas, un flux intra ou inter comtés, des mobilités locales (Tobler, 1995 : 329). Le second groupe est celui des flèches longues qui illustrent un mouvement inter-îles. Ces groupes peuvent ainsi être appréhendés comme l'illustration des échelles de la mobilité interne du Royaume-Uni : ils collent plus ou moins parfaitement avec la typologie proposée dans la section du texte qui précède la carte. La présence sur la carte de E.G. Ravenstein de ces différents groupes de flèches illustre, à notre sens, deux approches méthodologiques distinctes de la cartographie de flux, témoin d'un basculement entre une approche territoriale de la cartographie de flux fondée sur un maillage discret de l'espace et une approche spatiale fondée sur une acception continue de l'espace, mieux à même de représenter la dynamique des mouvements.

¹¹⁵ E.G. Ravenstein publie deux textes complémentaires sur ses lois, en 1885 et en 1889 qui feront date, eu égard aux discussions que le premier suscita (Greenwood et al., 2003).

Figure 4-7. La première carte de flux où figurent des flèches (d'après¹¹⁶ Ravenstein, 1885)



La matrice d'origine serait de type (Fij) ; les données, issues d'un RP, décriraient le solde bilatéral de ces navettes britanniques inter comtés.

¹¹⁶ La carte originale est d'une qualité insuffisante pour pouvoir être présentée en l'état dans la publication que nous lui avons consacrée (Bahoken, 2013a). C'est pourquoi nous l'avons entièrement redessinée en tentant de respecter, autant que possible, l'original.

Sur le plan méthodologique, la représentation des mouvements des migrants à l'aide de flèches, telle qu'elle est réalisée par E.G. Ravenstein, rompt avec les pratiques de ses contemporains et avec les nôtres. Si les cartes des flux immatériels suivent généralement le modèle dit des « atlas » (Müller, 1991), le procédé de E.G. Ravenstein semble suivre un tout autre modèle.

Dans la mesure où E.G. Ravenstein a eu accès à des données censitaires, nous faisons l'hypothèse qu'il a très probablement représenté ses courants migratoires selon une approche discrète qui s'appuie sur le maillage territorial irrégulier des comtés. Cependant, le fait que l'on distingue, par endroits, un effet local de tourbillon (Rosental, 1997) suggère une vision plutôt continue des mouvements et une représentation cartographique du type des champs vectoriels, tels qu'ils sont utilisés en physique. L'auteur n'a peut-être pas tracé des flèches exactes au niveau individuel, il aurait plutôt cherché à retranscrire une forme générale du mouvement correspondant plus à un sondage, qu'à une cartographie exhaustive. C'est, en effet, comme si l'auteur avait eu recours à des concepts utilisés en physique pour représenter ses flux migratoires.

4.1.3.1 *Le recours à la physique*

La méthode de représentation cartographique mise en œuvre par E.G. Ravenstein tient de l'approche discrète de la cartographie de flux. L'auteur s'est, en effet, appliqué à tracer de nombreuses petites flèches comme s'il reliait des couples de lieux situés à une échelle infra-comté. Cependant, le maillage des comtés semble, à première vue, ne servir qu'à l'habillage de la carte : il n'y a effectivement aucun caractère systématique dans le placement des flèches. L'auteur aurait, en réalité, été influencé par les travaux de ses contemporains physiciens pour représenter un phénomène discret de manière apparemment continue. La carte fait, en effet, directement écho au caractère physique de la dynamique du mouvement. Nous faisons l'hypothèse qu'elle témoigne très probablement de la première tentative de transfert des notions de champs vectoriels de pression depuis la physique vers la cartographie statistique, pour invalider l'affirmation de W. Farr. L'idée serait ainsi à l'origine d'une recherche de Lois sur la Migration.

4.1.3.2 *L'astuce graphique*

La cartographie de flux est traditionnellement réalisée selon le modèle graphique de la flèche, présenté dans la seconde section de ce chapitre. Dans la majeure partie des cas, les flux sélectionnés pour la représentation mettent en relation les centres géographiques des zones d'origine et de destination. Ils ont une forme droite ou curviligne pour tenter de limiter les chevauchements. Les principes de sémiologie graphique (Bertin, 1967) imposent le respect d'un certain nombre de règles de représentation, notamment la figuration d'un effet de convergence ou de divergence des flux, depuis ou vers les centres des zones.

Ainsi, indépendamment de leur forme, le tracé des liens a pour objectif d'illustrer le déplacement, par un franchissement des limites des zones, en reliant habituellement leurs centres géométriques. Si cette pratique est aujourd'hui courante, elle n'est pas celle E.G. Ravenstein. En effet, cet auteur n'a représenté que l'endroit du franchissement de la limite par des courtes flèches et non l'intégralité du lien entre les couples de lieux. Les multiples petites flèches illustrent ainsi davantage les points de passage des mouvements plutôt que la liaison entre les lieux. L'auteur a, en outre, systématiquement appliqué des angles au tracé de ses flèches afin de représenter de manière dynamique le mouvement. A noter que ce recours au procédé de la petite flèche courbe lui permet de marquer le flux, ainsi que sa direction, tout en évitant l'épineux problème du chevauchement des liens, en raison de *l'effet spaghetti* présenté dans le chapitre 2.

Toujours sur le plan graphique, la carte de E.G. Ravenstein a très certainement fait l'objet d'une simplification afin d'assurer la lisibilité des petits et des grands flux. Le problème de la longueur des

flèches s'est très probablement posé dans la réalisation de cette carte même si l'auteur n'en fait pas directement état. Il a, en effet, soulevé au début de son texte les lacunes des données et les contraintes qu'elles ont impliquées dans la réalisation de ses cartes, notamment la prise en compte de la distance qui sépare les lieux. Ainsi, même s'il n'est pas directement question de la carte n°5, l'auteur a mentionné la difficulté engendrée par l'hétérogénéité du maillage des comtés. Les variations de tailles des zones donnent ainsi l'impression « [...] *qu'un migrant traverse la frontière d'un comté avec un mouvement de 25 miles dans un cas, tandis que dans un autre, un mouvement de 95 miles peut ne pas franchir la limite d'un comté* » La distance¹¹⁷ entre les lieux, qui traduit la longueur du mouvement, fut effectivement l'une des notions auquel il s'intéressa pour l'élaboration de sa seconde Loi sur les migrations. Nous supposons qu'il en a été de même pour la réalisation de sa carte des Courants migratoires. Cependant, E. G. Ravenstein a très probablement souhaité rendre compte de l'effet de la distance sur les courants migratoires, de manière graphique, mais sûrement pas en usant d'artifices. Cette hypothèse justifie très certainement le recours aux petites flèches courbes, mais pas seulement.

La pratique courante actuelle de cartographie de flux consiste à relier les centres des zones d'origine et de destination. Cette commodité privilégie visuellement les zones de grande taille qui se voient donc affublées d'un long lien prenant une place importante. Il s'agit là de l'effet d'absorption visuelle des flux par les zones de grande taille, déjà mentionné dans le chapitre 2. Ce principe général assimile l'ensemble de la population des zones d'origine et de destination à leur barycentre, ce qui n'est pas le cas dans la réalité : un centre récepteur ou émetteur peut être théoriquement situé en tous points de la zone concernée. Cette pratique introduit aussi, à notre sens, une distorsion générale dans la perception de la distance parcourue par le flux, tendant à favoriser les grandes zones au détriment des plus petites. Elle complique également l'appréhension de la structure des flux par la variation des *longueurs* des liens par rapport à la distance qui sépare les centres des couples de zones.

Pour éviter les écueils liés à la prise en compte de la distance sur un maillage hétérogène, nous supposons que E.G. Ravenstein a souhaité traiter les couples de zones contigües – il s'agit là de migrations locales – de la même façon, quelle que soit leur taille. C'est la raison pour laquelle l'auteur a préféré recourir au placement de courtes flèches courbes aux seuls points de franchissement des limites inter zones, indiquant la tendance du mouvement. De la sorte, les petites flèches courbes représentent, à notre sens, des flux transfrontaliers locaux qui se sont produits soit entre comtés, soit entre des localités proches. L'auteur a ainsi réservé l'utilisation des longues flèches, comme nous l'avons vu, aux flux inter îles qui représentent des trajectoires particulières. Concernant le tracé des longues flèches, E.G. Ravenstein a utilisé le modèle de la trajectoire interrompue en différents points centraux plutôt que multiplier les petites flèches. Cette astuce lui permet de renseigner l'observateur sur la présence éventuelle d'étapes dans le déplacement et sur la catégorie de migrants concernés identifiée dans sa typologie, sans surcharger la figure. Au-delà de la trajectoire, il s'agit là probablement de la mise en œuvre du procédé de C.-J. Minard – sans variation de taille – où les liens les plus importants se scindent en plusieurs bras. Ces grandes flèches illustrent des flux qui se produisent effectivement sur de très longues distances, vers des centres industriels et commerciaux en illustration de l'un des déterminants de ses courants migratoires.

E.G. Ravenstein aurait donc adopté une approche territoriale de la cartographie de flux, tout en donnant l'impression d'une approche spatiale pour représenter la continuité et la dynamique des mouvements. Pour ce faire, il a eu recours à une astuce qui consiste à ne tracer que le franchissement

¹¹⁷ Outre la distance entre les lieux, les deux autres notions mentionnées dans la seconde Loi des migrations de Ravenstein sont liées à l'attractivité des localités au regard de leur dynamisme économique ainsi qu'à la dépendance des flux par rapport à la masse des lieux d'origine et de destination. Ces trois notions, couplées à celles des deuxième et troisième lois sur le caractère push-pull des migrations entre les espaces ruraux et urbains seront à l'origine de la formulation gravitaire des modèles d'interaction spatiale (Ravenstein, 1885, pp. 186-187).

des limites de zones. C'est en ce sens que nous considérons que la méthode mise en œuvre dans la carte n°5 de Ravenstein peut contribuer à renouveler la pratique actuelle de cartographie de flux.

L'apport de E.G. Ravenstein à la cartographie de flux est indéniable pour une raison majeure : à l'époque où l'auteur réalise sa carte, l'analogie des mouvements – ou des flux spatialisés, comme nous le verrons dans le chapitre 6 – avec les lois de la gravitation n'a pas encore été formalisée. On sait par ailleurs que le transfert de notions physiques à la cartographie de flux qu'il a opéré s'est révélé particulièrement fécond puisqu'il permettra, quelques années plus tard, notamment à C. W. Thornthwaite¹¹⁸ (1934), cité par (Tobler, 1994, 1995 ; Wheeler, 2002), puis à W. Tobler (1987) de les mettre en pratique pour réaliser des cartes de flux qui feront ensuite date

L'une des propositions méthodologique de D. Dorigo et W. Tobler (1983) est en ce sens symptomatique de cette période. Les cartes de flux migratoires suisses que ces auteurs proposent s'appuient sur la notion de « *champs vectoriels* » issue de celle de potentiel¹¹⁹. La Figure 4-8 illustre ces cartes à deux niveaux de résolution spatiale¹²⁰ : la carte de gauche est réalisée sur un maillage spatial en 184 districts (14,7 km de résolution), et celle de droite porte sur 26 cantons (39,2 km de résolution). Le procédé graphique de représentation est celui des champs de potentiels où les flèches représentent des iso-potentiels.

Figure 4-8. Illustration de l'anamorphose vectorielle multipolaire



Source : D. Dorigo et W. Tobler, 1983 ; W. Tobler (1990).

Simultanément, à cette période qui précède la période dite du « tournant spatial », s'opère une autre révolution dans la manière de penser et d'analyser les flux : un tournant global. A la faveur d'un changement d'échelle d'observation, la vision régionale commune est momentanément substituée à une vision mondiale des échanges, comme un décollage lié au développement de l'imagerie aérienne. Cette posture traduit une nouveauté dans la manière de penser les réseaux et les flux qui conduiront, quelques décennies plus tard, à l'émergence de la notion de mondialisation.

Sur le plan cartographique, l'émergence d'une vision mondiale conduit à une nouveauté : la cartographie de relations (Lij) plus que de valeurs de flux (Fij), à une échelle globale et non plus seulement locale. Cette nouvelle vision globale de la cartographie de flux s'inscrit dans ce que des

¹¹⁸ Thornthwaite, C., W., 1934. *Internal Migration in the United States*, Dissertation, University of Pennsylvania, Philadelphia, cité par W. R. Tobler (1994, 1995).

¹¹⁹ Cette utilisation du potentiel s'appuie sur une hypothèse forte de la distribution des phénomènes spatiaux selon laquelle l'attractivité d'un lieu est proportionnelle à sa masse et inversement proportionnelle à la distance qui le sépare des autres lieux (attirés).

¹²⁰ La résolution spatiale fait référence à la taille de la maille territoriale et par conséquent, à l'échelle géographique d'observation du phénomène.

géographes contemporains tels que E. de Martonne, Membre de l'Institut, qualifient d'ère de la « géographie aérienne ».

4.1.4 La vision globale (E. Raisz ; J. Bartholomew, 1920- ; Renner, 1942), l'ère de « la géographie aérienne »

La réalisation de cartes de flux à l'échelle mondiale émerge puis se développe de manière fulgurante dès les années 1920-1930. Elle donne naissance à un domaine foisonnant qui consiste à penser, à analyser et à cartographier les réseaux et les flux à l'échelle globale de la planète terrestre. Si les représentations des flux commerciaux internationaux réalisées par C.-J. Minard figurent les premières cartes statistiques, les astuces de généralisation des contours qu'il utilise volontairement, notamment la déformation des traits de côtes pour faciliter le placement des figurés au droit des zones d'origine et de destination, en font des figures davantage artistiques que formellement exactes, sur le plan géométrique. Le contexte de la deuxième guerre mondiale, les progrès réalisés dans le domaine de l'aviation, notamment les possibilités de prises de vue aériennes et, plus loin, celles de la cartographie aérienne, ont entraîné un renouvellement important de la cartographie de flux et des réseaux. Celle-ci est désormais globale : elle prend en compte la spécificité de la sphère qui se retrouve dans le rendu.

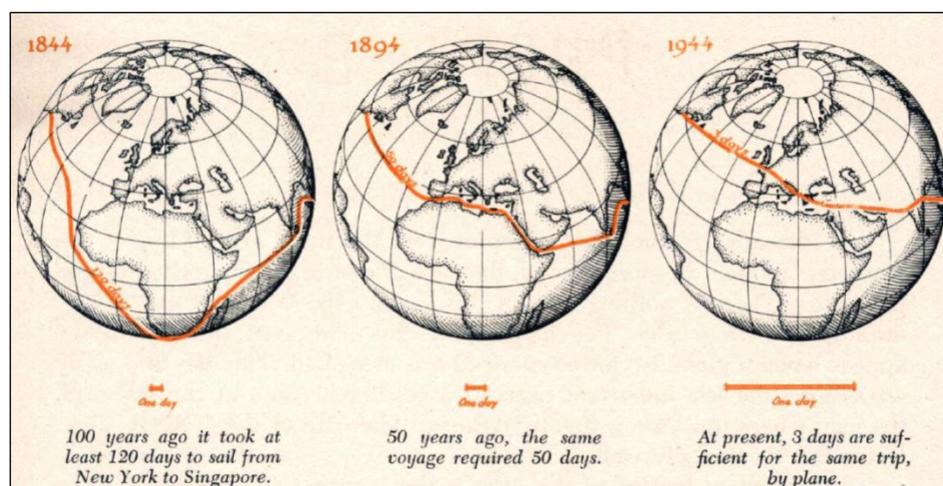
C'est la volonté de compréhension des enjeux stratégiques et surtout pacifiques¹²¹ liés à la deuxième guerre mondiale à une échelle globale, plutôt que régionale ou locale, associée à la nécessité d'une description plus précise du rétrécissement du monde lié aux progrès de l'aviation, et des forces géopolitiques qui en découlent, qui ont ouvert l'ère aérienne de la géographie humaine (Renner, 1942 ; Raisz, 1944 ; Engelhard, 1943). L'émergence de cette vision globale a donné lieu à l'élaboration de nouveaux types de documents cartographiques qui sont magnifiés dans des ouvrages symptomatiques parmi lesquels *l'Atlas Fortune* (Fortune 1941), *l'Atlas of global Geography* (Raisz, 1944) ou encore l'ouvrage *Human geography in the Air-Age* (Renner, 1942).

G.T. Renner présente en 1942 un ouvrage qui est révélateur de l'importance de la production cartographique en manuels, en lithographies et atlas à l'échelle mondiale. L'ouvrage caractérise cette « ère aérienne », liée à l'accroissement des vitesses de déplacement par voie aérienne entraînant une modification de la vision du monde liée à l'amélioration de l'accessibilité des villes connectées, caractéristique d'un rétrécissement du monde (voir Figure 4-9).

Cette modification de la représentation de la terre s'inscrit dans le contexte de la réouverture d'une discussion ancienne entamée par les grecs, puis abandonnée, sur le choix de la projection cartographique. La vision polaire, en deux voire en trois dimensions, a ainsi supplanté la vision rectangulaire du monde de G. Mercator (1569) chez les auteurs emblématiques de la période. E. Raisz (1893-1968), par exemple, géographe et cartographe Hongrois et Américain de l'Université d'Harvard, fait partie de ces auteurs qui ont fortement contribué au renouvellement de la cartographie. D'après A. Robinson (1970), E. Raisz a été présent au bon moment et au bon endroit. « *Qui mieux que lui pouvait savoir ce qu'est une carte mondiale* » ? rappelle l'auteur dans l'introduction de l'article qu'il lui consacre.

¹²¹ Nous considérons que l'ère de la géographie humaine aérienne a fortement contribué à la paix mondiale, alors que les enjeux de son émergence s'inscrivent dans le contexte de la seconde guerre mondiale et pour cause, la géographie ayant surtout servi, d'abord, à faire la guerre.

Figure 4-9. L'ère aérienne et le rétrécissement du monde



Source : Eisler G., B., 1944, "Introduction" in E. Raisz, 1944, *Atlas of Global Geography*.

E. Raisz est un auteur doté d'un véritable génie cartographique pour plusieurs raisons. Il publie en 1938, le premier manuel de cartographie¹²² en langue anglaise, à une époque où les enseignements étaient davantage fondés sur la pratique et l'expérience (Robinson, 1970 : 1). L'*Atlas of Global Geography* qu'il réalise en 1944 est dédié à l'aviation alors que l'auteur était surtout connu jusque là pour ses travaux en géographie physique. Il s'intéresse ainsi de manière avant-gardiste à la cartographie des relations internationales rendues possibles par le développement de l'aviation.

La planche de la Figure 4-10 justement, intitulée *The Air Age*, est extraite de cet atlas. Réalisée à main levée, à l'aide d'un stylo et d'une plume à encre, elle fait partie des figures représentatives des travaux de l'auteur sur la période.

L'originalité de cette figure réside dans le choix du fond de carte et de sa projection cartographique sur lesquels nous reviendrons dans le chapitre 6¹²³, aussi de la thématique représentée. Le fond de carte est celui des terres connues à l'époque, représentées sur un globe, en trois dimensions. L'auteur semble avoir souhaité représenter la rotondité de la terre de manière non conventionnelle pour l'époque, caractérisée par des cartes polaires souvent planes, c'est-à-dire représentées en deux dimensions et non sous la forme d'un globe.

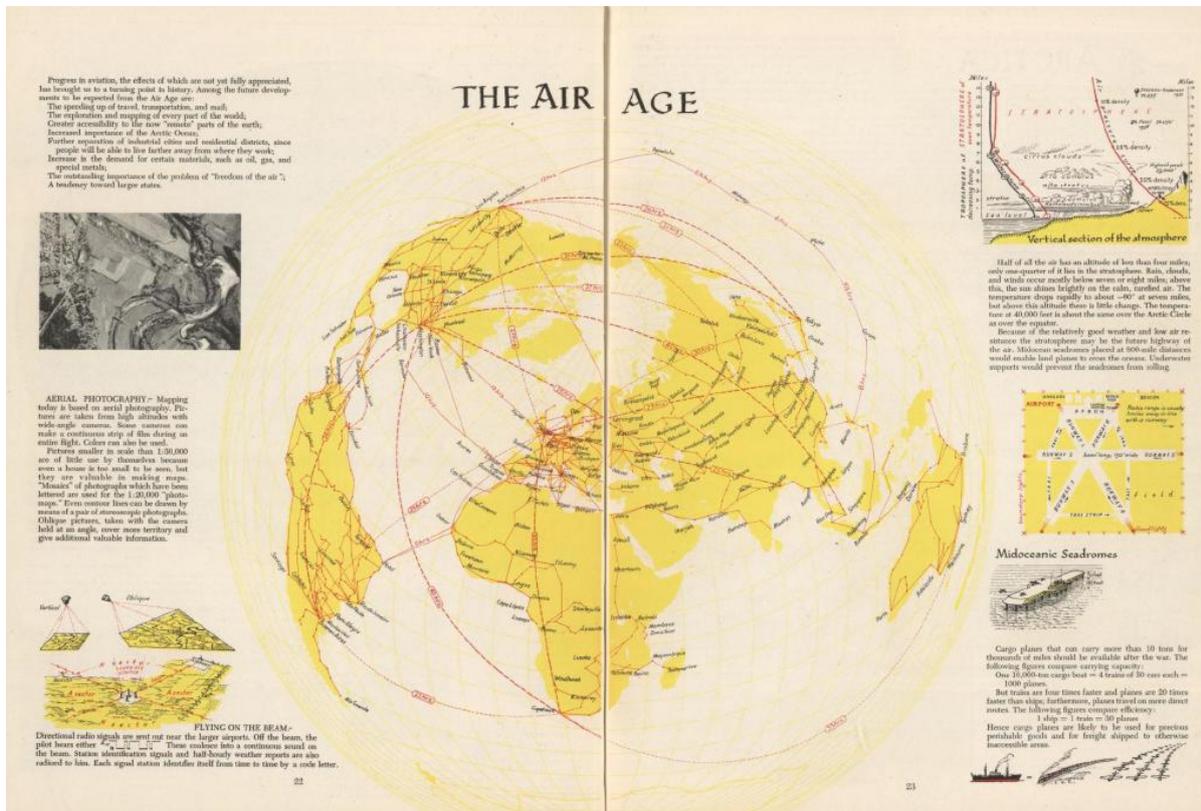
Le génie cartographique de E. Raisz tient, de notre point de vue, dans le recours à la troisième dimension et cela, plus de soixante ans avant le projet *Earth* de Google©. La vision du monde sur un globe permet à l'auteur de privilégier la représentation d'un monde en réseau, doté de plusieurs centralités, là où ses contemporains présentent des cartes mondiales mono centrées : sur les Etats-Unis, l'Australie, l'Europe ou encore l'Afrique (Harrison, 1943 : 52-53), comme dans *l'Atlas Fortune* (Harrison, 1941). Cette représentation d'un monde en réseau souligne l'importance des relations inter urbaines à l'échelle mondiale, mais aussi des trajectoires commerciales aériennes, comme sur la carte de la Figure 4-10, ou des échanges maritimes situés plus loin dans l'ouvrage (Raisz, 1944 : 52-53). Il

¹²² Raisz, E., J., 1938, 1948, *General cartography*, McGraw-Hill Series in Geography, New York, London.

¹²³ Le choix de systèmes de projections cartographiques polaires, ou de représentations sous la forme de globe, est une réponse apportée à un problème spécifique de cartographie des flux et des réseaux à l'échelle mondiale. Il correspond à un *effet d'alignement topologique* (ou de position des lieux), nous y reviendrons dans le chapitre 6.

faut dire que l'époque connaît les premières typologies d'un système de villes mondiales en réseaux qui soulignent leur spécialisation : résidentielle ou industrielle.

Figure 4-10. Une représentation globale des flux et réseaux, selon E. Raisz (1944)

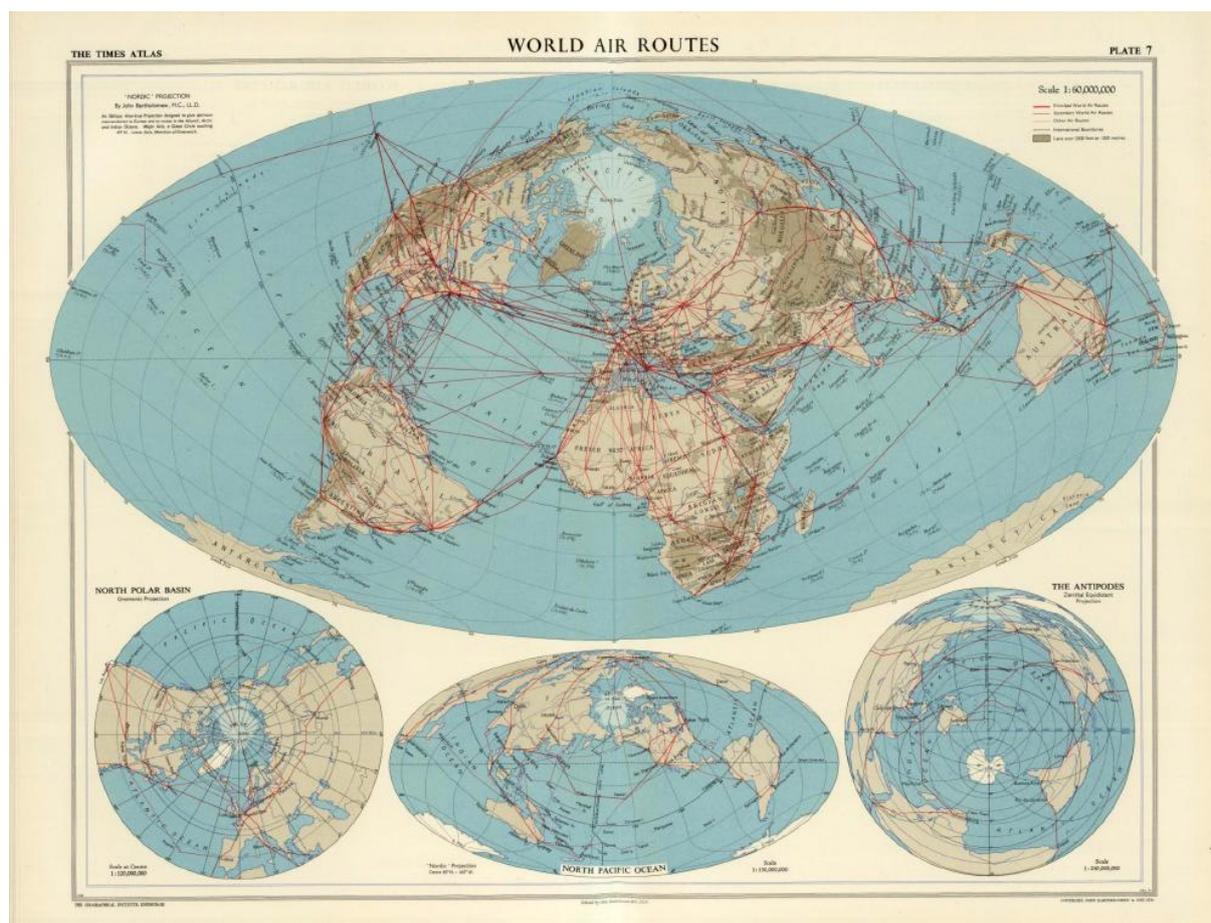


Source : E. Raisz, *Atlas of Global Geography*, 1944.

La matrice de flux utilisée pour réaliser cette planche est nécessairement dégradée (voir chapitre 3) puisque seule l'existence d'un lien (Lij), probablement d'une relation bilatérale, est représentée par ces lignes. Celles-ci sont courbes, non parce que l'auteur a eu une velléité esthétique mais parce qu'elles suivent la courbure de la sphère : la métrique utilisée est, à juste titre, orthodromique et non euclidienne. Ce qui semble primer pour l'auteur c'est davantage la mise en relation des lieux – par ces liens- à une échelle globale, mondiale, que la visualisation de quantités échangées à l'échelle locale ou selon une vision mono centrée sur une puissance. Cela a d'autant plus d'importance que l'auteur a pu réaliser des cartes du monde, dans le même atlas, du point de vue des seuls Etats-Unis.

La représentation globale du monde a également eu pour conséquence d'accroître la visibilité des contrées reculées ou récemment découvertes. C'est ainsi que la Nouvelle Guinée et l'Arctique, par exemple, peu connues à l'époque se retrouvent placées au centre de débats liés à la prise de conscience de prometteuses voies de circulation internationales. Elles seraient d'ailleurs représentées pour la première fois sur la planche de la Figure 4-11, située page suivante.

Figure 4-11. Les routes aériennes mondiales



Bartolomew, J.-L. *Time Atlas of the World* (1958)

La matrice de flux utilisée est identique à la précédente : il s'agit d'une matrice d'adjacence (L_{ij}) caractéristique de cette volonté de représenter le réseau de relations (de villes) mondiales issue du développement de l'aviation ; les lignes décrivant en fait les principales voies aériennes. La volonté d'exhaustivité de l'auteur explique probablement les multiples points de vue proposés, selon différentes projection cartographique, et leur disposition sur une même planche.

La planche ci-dessus est à l'origine une lithographie de 431 * 557 mm reproduite dans le *Time Atlas of the World*¹²⁴, réalisé entre 1955 et 1959. Cet ouvrage conséquent est consacré aux changements géopolitiques survenus à la suite de la seconde guerre mondiale : il est composé de 129 planches réparties en cinq volumes, chacun étant consacré à un continent.

La planche reprise ici fait partie du premier volume intitulé *The world Australasia and East Asia*. Elle décrit les routes aériennes mondiales selon des projections polaires centrées sur différentes régions :

¹²⁴ La série d'atlas éponyme a été impulsée par J. L. Bartholomew en 1885, un célèbre cartographe du *Geographical Institute of Edimburg*, médaillé d'or de la *Royal Scottish Geographic Society*. L'auteur est issu d'une véritable dynastie de cartographes puisqu'il est le sixième du nom. La *National Library of Edimburg* en Ecosse leur a d'ailleurs consacré une très belle exposition photographique en 2013, intitulée *The World of Jon Bartholomew and Sons* dont une partie est visible en ligne, à l'adresse suivante (<http://johnbartholomew.com/Meet%20the%20Bartholomews%20Exhibition/>) [vérifiée le 8 septembre 2013]. A noter que l'auteur est aussi connu pour avoir initié le principe de la coloration continue des cartes en isolignes.

l'Europe occidentale correspond à la carte principale, elle est centrée à 45° de latitude nord ; les autres cartes sont centrées sur le pôle nord, sur la région pacifique nord et sur les Antipodes (ou pôle sud).

Ces représentations des routes mondiales placent l'océan Arctique comme une voie d'accès permettant de relier rapidement l'Asie à l'Europe (Raisz, 1944 : 50-51). La production cartographique s'oriente alors vers le tracé de routes, ainsi que dans la figuration des temps de déplacement en fonction du mode de transport concerné.

E. Raisz a, par exemple, présenté une courte comparaison de l'efficacité du transport aérien de marchandises par rapport aux autres modes de transport que sont les voies ferroviaires et maritimes, en ces termes.

« *Un avion-cargo de 10 000 tonnes = 4 trains de 50 wagons chacun = 1000 avions. Mais les trains sont quatre fois plus rapides et les avions vingt fois plus que les cargos ; de plus, le transport en avion emprunte des routes plus directes. Les figures suivantes en comparent l'efficacité : 1 cargo = 1 train = 30 avions. L'avion-cargo est, par conséquent, susceptible d'être utilisé pour les marchandises périssables ou précieuses et pour l'expédition de fret dans les zones inaccessibles* » (Raisz, 1944).

La représentation d'un monde en réseau met en évidence sa connexité, ainsi que les différentes routes permettant d'aller d'un point à un autre. Elle permet aussi de révéler des contrées jusqu'ici méconnues, de mettre en évidence le rôle des relations aériennes dans la définition de régions mondiales.

Dès lors, la popularité de ces fonds de cartes mondiales n'a eu de cesse de croître notamment en réaction à la vision de Mercator. Deux projections mondiales apparaissent aujourd'hui populaires, la projection de Gall-Peters (1973) supposée représenter de manière équitable la répartition des masses continentales et la *Dymaxion map*¹²⁵, ce planisphère de Buckminster Fuller (1954) souvent considéré comme le mieux adapté à la représentation des flux à l'échelle globale.

Les discussions sur les avantages comparés de différents systèmes de projection intéressent surtout la cartographie des processus qui s'expriment à l'échelle mondiale, le choix de l'un ou de l'autre système n'étant pas de nature à modifier les cartes réalisées à grande échelle. Pour celles-ci, les débats de la première moitié du XX^{ème} siècle portent davantage sur la manière de mieux représenter les flux et pour cela, font appel à des principes issus de la physique qui conduiront, de notre point de vue, à l'apogée de la cartographie de flux.

4.1.5 L'apogée (C. W. Thornthwaite, T. Hägerstrand, W. Tobler, 1934-1987)

L'intérêt du recours à des méthodes physiques pour représenter les flux tient dans la visualisation de manière discrète, de la dynamique d'un processus observé empiriquement. Si l'innovation apportée par E.G. Ravenstein, qui consiste à représenter des champs migratoires à la manière de champs de vecteur est essentiellement d'ordre graphique, les auteurs qui s'en sont inspirés ont formalisé mathématiquement cette idée, en transférant des lois physiques vers la géographie humaine pour expliquer la répartition spatiale de phénomènes sociaux.

4.1.5.1 Le transfert de lois physiques vers la géographie

Les travaux de C.W. Thornthwaite (1934), cité par W. Tobler (1995), peuvent être considérés comme le point de départ d'une nouvelle démarche qui s'appuie sur des principes physiques pour expliquer des phénomènes sociaux (voir Encart 4-2).

¹²⁵ *Dymaxion : Dynamic Maximum Tension.*

Encart 4-2. De la physique à la géographie des flux

« En Amérique, comme ailleurs, la migration est un processus qui dépend des moyens de communication entre des zones qui ont différentes intensités de pression de population. Ces gradients de pression sont amenés à croître dans une zone et à décroître dans une autre. L'intensité relative de la pression de population peut être croissante pour une zone donnée en fonction d'une 'contraction' des opportunités économiques et sociales ou par la croissance continue de la population, elle peut être décroissante en raison d'une extension des opportunités ou du fait d'une baisse de la population. Par les flux de population provenant de régions qui présentent une haute pression importante, vers des régions qui présentent des basses pressions, les inégalités tendent à se réduire. Ce qui est important dans les migrations est l'inverse de la relation de résistance, à la fois physique et culturelle, qui est rencontrée. L'isolement physique, l'inertie, le préjudice, et l'ignorance sont autant de facteurs qui inhibent plus ou moins la liberté des mouvements de population. Les flux de population sont ainsi, en quelque sorte, analogues aux flux d'électricité, la formulation mathématique qui en découle pour les migrations mérite une précision* [% suivant]. Le total des migrations d'une zone à une autre est directement proportionnel à la pression du gradient entre elles, et inversement proportionnel à la résistance » (Thornwaite, 1934 : 1).

C.W. Thornwaite a porté en note de bas page de son texte, traduit par nos soins, la formulation mathématique des gradients de pression adaptée aux migrations de population, en référence à loi d'Ohm.

W. Tobler (1995) précise dans un texte consacré aux travaux de C.W. Thornwaite et E.G. Ravenstein que, si le premier situe son propos dans le champ scientifique de la physique, c'est certainement en raison de son attrait pour les sciences de la terre et l'économie. Les notions mobilisées sont ainsi celles des analyses coût-bénéfice : le texte de C.W. Thornwaite ne fait donc pas référence aux concepts d'interaction spatiale ou de modèle gravitaire. L'auteur n'est pourtant pas connu pour avoir travaillé sur des phénomènes économiques spatialisés mais plutôt sur des processus physiques, en climatologie et en hydrologie.

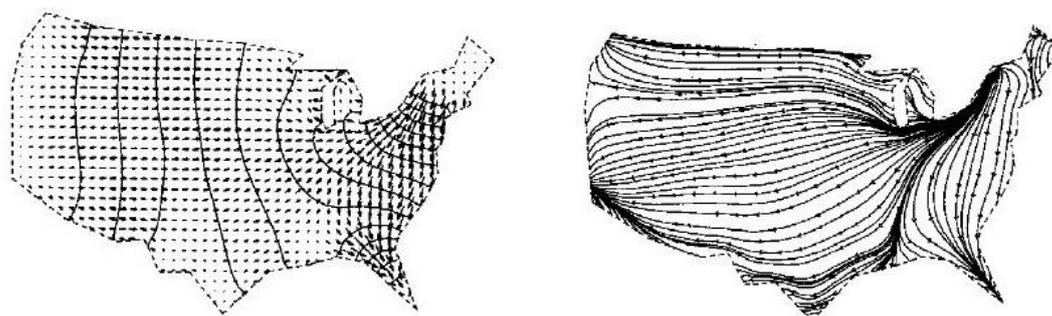
La monographie de 52 pages que de C.W. Thornwaite réalise au début de sa carrière d'enseignant, à l'Université d'Oklahoma, « [...] mérite un examen attentif car la cartographie relève du challenge » (Tobler, 1995 : 329). Elle présente, en effet, de nombreuses cartes innovantes qui portent notamment sur des flux migratoires et font écho aux notions de physique électrique de champs de pression, de gradients et de potentiel. Quelques décennies plus tard, l'économiste Lucas (1981) mentionné par W. Tobler (1995) reprendra les fondements de cette recherche en comparant, à son tour, la migration à un courant électrique, de la manière suivante : « [...] un ajustement du flux répond à des différentiels de pression qui se manifestent à l'opposé, à la fin du pipeline. Cette idée suggère que ce n'est pas le niveau absolu de pression ni de poussée qui est en cause, mais l'existence d'une relative attraction entre les éléments ». (Lucas, 1981 : 85, cité par Tobler, 1995 : 329).

C'est ainsi à la suite de C.W. Thornwaite et de T. Hägerstrand que W. Tobler et D. Dorling, notamment, ont développé des recherches qui les ont conduits à proposer des méthodes de cartographie de flux résolument innovantes, dont la complexité est, pour certaines, toujours d'actualité. On trouve parmi celles-là, la cartographie en anamorphose mais également l'analogie avec les champs de pression en physique, sous la forme de champ migratoires.

4.1.5.2 Les champs migratoires

W. Tobler indique repartir de l'idée de C.W. Thornwaite (1934) de recourir aux notions physiques de pressions et de gradients pour cartographier des flux. L'utilisation des champs de potentiel, comme méthode de traitement puis de représentation cartographique des flux dès les années 1930, constitue une véritable avancée méthodologique. L'exemple typique d'application peut être illustré par la série de cartes des flux financiers entre les Etats américains présentée sur la Figure 4-12, elles ont été réalisées par D. Dorigo et W. Tobler (1983) et W. Tobler (1981, 1981-c,) respectivement statisticien et géographe-cartographe, ce qui n'est pas sans importance sur les résultats obtenus.

Figure 4-12. Les champs migratoires



Source : D. Dorigo, W. Tobler, 1983.

Ces cartes ont été réalisées à partir d'une matrice antisymétrique de type (Fij), Elles décrivent les directions préférentielles des transferts nets, c'est-à-dire la partie dynamique de la matrice que l'auteur aurait très probablement souhaité représenter à l'aide d'isolignes courbes. Ces cartes correspondent à l'exemple typique de représentation de cette composante des flux par l'auteur.

W. Tobler indique, en 1983, qu'il s'est fondé (toujours avec l'aide de D. Dorigo), sur les principes de la loi d'Ohm à l'instar de l'économiste Lucas pour modéliser les flux au préalable de leur cartographie. Les auteurs considèrent ainsi que l'intensité de la migration est proportionnelle à la pression et inversement proportionnelle à la résistance, ce qui leur permet d'introduire des analogies avec les principes newtoniens d'attractivité des lieux et de distance. Les deux auteurs en arrivent à la conclusion selon laquelle les flux donnent lieu à un champ de potentiel qui satisfait une équation de Poisson.

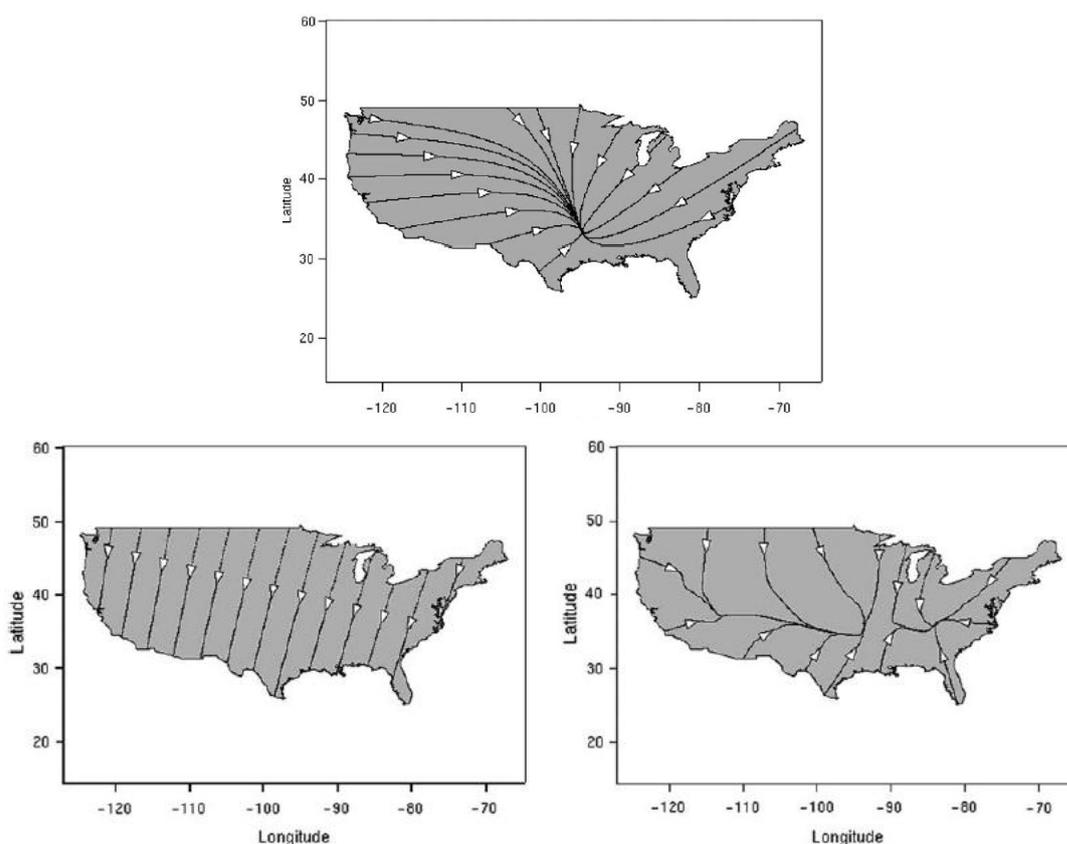
La Figure 4-12 est le résultat obtenu suite à la réalisation de plusieurs traitements de données réalisés par le binôme Dorigo - Tobler qui ont conduit à différentes cartes considérées comme intermédiaires. La première étape a consisté à appliquer la méthode des quadra, caractérisée par une transformation du maillage hétérogène des 48 Etats américains en un maillage régulier rectangulaire. L'application d'une grille régulière (de 91*65 nœuds) sur le maillage hétérogène des comtés permet le dénombrement du nombre de liens qui franchit les différentes limites de mailles, et le décompte du total des flux associé à chaque lien. Les auteurs ont fait le choix de calculer un solde bilatéral inter maille qui les met en présence d'une matrice de type (Fij). La première carte qui en résulte est formée de courbes d'isopotentiels : elle constitue la *surface de potentiel*. La seconde carte porte sur des *champs de vecteurs* représentés sous la forme de gradients en quatre classes de valeurs. Ces classes représentent des intensités croissantes ou décroissantes d'échanges entre les lieux. La combinaison de ces deux cartes (de potentiel et de champs de vecteur) conduit à une troisième carte : celle des directions préférentielles des échanges (voir Figure 4-12) qui présente les vecteurs-résultats des *soldes*

migratoires *bilatéraux* entre les lieux, indiquant « [...] *la principale direction des échanges (transferts nets)* » (Cattan, Grasland, 1994 : 116).

L'idée de représenter la direction préférentielle des échanges a aussi été appliquée par d'autres auteurs tels que C. Brundson et M. Charlon (2006), dont une série d'images des flux (Figure 4-13).

Les auteurs y combinent la méthode à un lissage, dans l'objectif de représenter la direction préférentielle locale des flux inter-états américains : « [...] *un modèle statistique probabiliste de vraisemblance locale plutôt qu'un modèle de champ de potentiel pour estimer l'orientation des flux* » (Brundson, Charlon, 2006 : 138) est utilisé. Ils calculent ainsi pour tout lieu d'origine (i), une probabilité directionnelle à partir de la moyenne pondérée localement des directions de l'ensemble des flux observés. Des isolignes d'égales directions sont ensuite représentées, elles illustrent la tendance des flux (voir Figure 4-13, en haut) qui, lorsqu'elle est combinée avec une méthode de lissage (voir Figure 4-13, en bas), permet de représenter les différentes échelles du système migratoire ce qui constitue, d'après les auteurs, l'un des apports de la méthode.

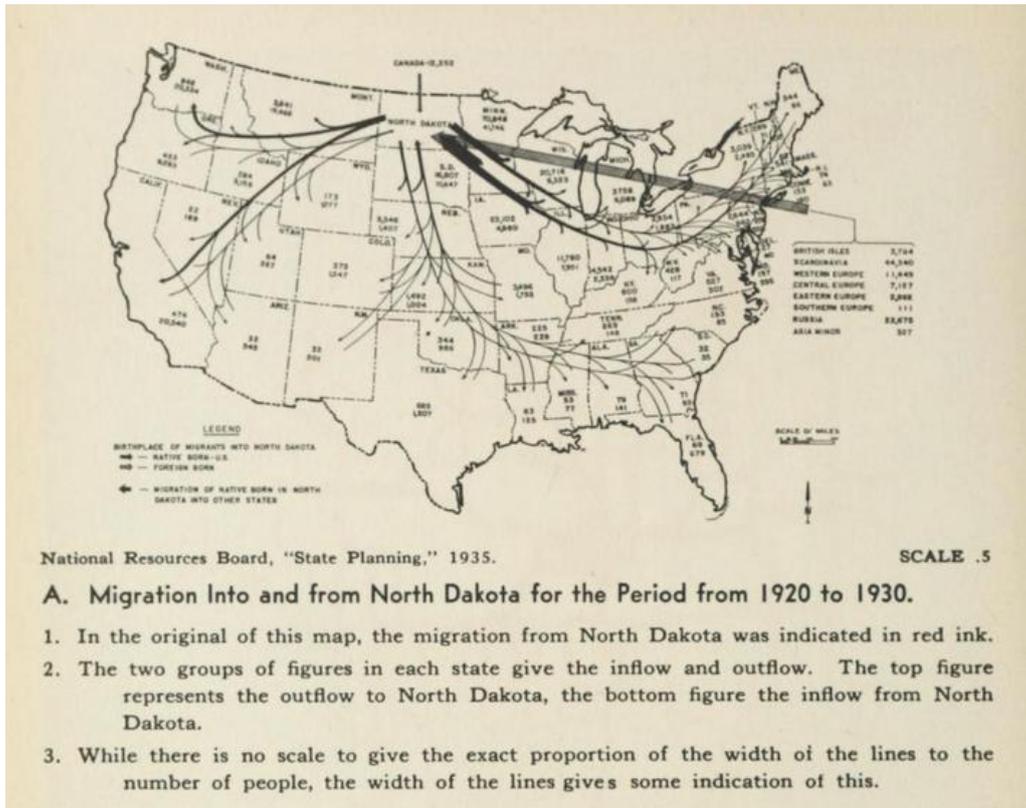
Figure 4-13. La direction préférentielle des flux



Source : C. Brunsdon, M. Charlon, 2006.

Plusieurs autres travaux de cartographie mettent en œuvre des procédés graphiques innovants, pour l'époque et pourraient l'être de même aujourd'hui, tels que ceux présentés sur la Figure 4-14.

Figure 4-14. Deux figures de la carte de flux de la première moitié du XXe siècle

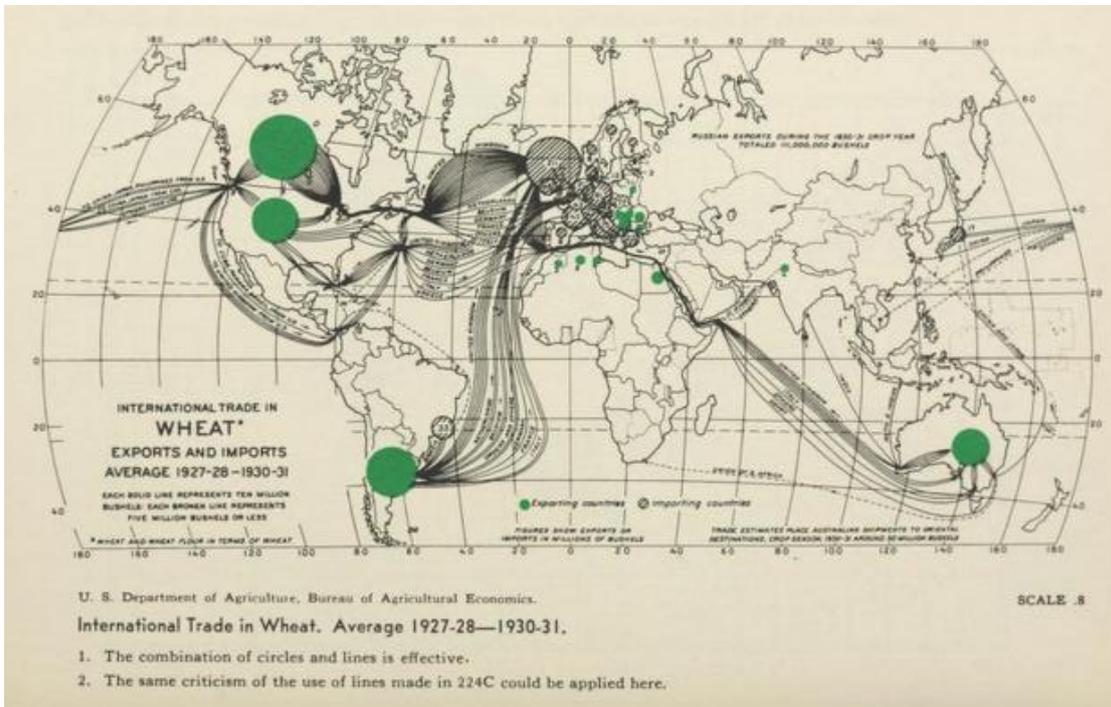


National Resources Board, "State Planning," 1935.

SCALE .5

A. Migration Into and from North Dakota for the Period from 1920 to 1930.

1. In the original of this map, the migration from North Dakota was indicated in red ink.
2. The two groups of figures in each state give the inflow and outflow. The top figure represents the outflow to North Dakota, the bottom figure the inflow from North Dakota.
3. While there is no scale to give the exact proportion of the width of the lines to the number of people, the width of the lines gives some indication of this.



U. S. Department of Agriculture, Bureau of Agricultural Economics.

SCALE .8

International Trade in Wheat. Average 1927-28—1930-31.

1. The combination of circles and lines is effective.
2. The same criticism of the use of lines made in 224C could be applied here.

Source: Brinton, W-P., (eds.) (1939).

Réalisés entre la fin du XIX^{ème} siècle jusqu'aux années 1930, ces travaux sont repris, pour l'essentiel, dans l'ouvrage *Graphic Presentation* paru en 1939, en l'honneur de W. Playfair. Ils ne semblent pas avoir fait l'objet d'une reprise, ou d'une actualisation.

Les nouveaux procédés de représentation des flux en champs de vecteurs, mis au point par D. Dorigo et W. Tobler (1983), sont liés à l'émergence d'une nouvelle manière d'appréhender l'espace et la matrice de flux.

4.1.5.3 Vers un nouveau paradigme de la cartographie de flux ?

Jusqu'ici, l'approche couramment mise en œuvre dans l'analyse cartographique des flux est discrète. Elle s'appuie sur des valeurs de flux mesurées entre des unités spatiales distinctes, non nécessairement contiguës, au sein de l'espace d'étude représenté par un découpage administratif. Lorsqu'elle est réalisée du point de vue des lieux, elle conduit à la réalisation d'une carte statistique en aplats, dite choroplèthe, selon le procédé mis en œuvre par C. Dupin (1784-1873), sachant qu'elle peut désormais être cognitive ou sociale (Dorling, 1991) et faire référence au principe de l'anamorphose.

L'approche discrète caractéristique de la carte choroplèthe est définie par J. Palsky (2012), dans un article de l'encyclopédie en ligne Hypergéométrie de la manière suivante. « *Il s'agit d'une représentation de quantités (plethos) relatives à des espaces, ou aires géographiques (khorê), par le moyen d'une échelle de tons gradués* »¹²⁶. Le procédé n'est pas dénué d'efficacité pour cartographier par différentes méthodes des résultats issus de dénombrements ou de mesures à des fins de classement, de hiérarchisation ou encore de comparaison dans le temps et/ou dans l'espace des valeurs prises par les différentes unités spatiales. Sa mise en œuvre revêt cependant une double complexité. Dans la l'encyclopédie en ligne Hypergéométrie, « *La carte choroplèthe pose un problème à la fois mathématique et graphique. Sa réalisation repose d'abord sur le choix d'une méthode de discrétisation, c'est-à-dire de division de la série statistique que l'on veut cartographier en classes, ou intervalles* » (*ibid.*)

Les méthodes de discrétisation sont appliquées sur des données qui sont nécessairement continues ou relatives, dans l'objectif de les découper en classes. Plusieurs méthodes de discrétisation des séries statistiques, au préalable de leur cartographie peuvent être mises en œuvre (Jenks, Coulson, 1963 ; Cauvin, Raymond, Serradj, 1987 ; Poitevin, 1999, Cauvin et al., 2008). Toutes conduisent à une partition de la série continue des données en classes assorties de bornes. Les données sont dès lors rangées (ou segmentées) en n classes plus faciles à manipuler. Elles sont cependant segmentées puisque ne sont dès lors connues que les informations qui appartiennent à des objets géographiques bien identifiés, appartenant à l'une des classes retenues.

L'approche continue permet de remédier à cet effet en conservant la couverture exhaustive de la série et de l'information sur l'espace d'étude et cela, sans introduire de découpage en classes. Il s'agit d'un changement de paradigme conceptuel rendu possible par le transfert des méthodes issues de la physique, des champs de pression à la géographie humaine. W. Tobler (1987) a, en effet, bien mis en évidence les conséquences d'une approche continue sur la cartographie de flux, notamment pour résoudre, graphiquement, le problème de *l'effet spaghetti*.

D'après Tobler (1987), une stratégie efficace de résolution de la complexité graphique de la carte de flux consiste à considérer que « [...] *la limite N des flux tend vers l'infini* » (Tobler, 1987 : 156). De ce point de vue, « [...] *l'espace géographique est considéré comme continu, et non pas divisé en unités statistiques de collecte. La table n'est plus limitée* » (Tobler, 1987 : 156). L'espace, ainsi assimilé à une surface continue est alors supposé contenir des valeurs en tous points et plus en certains. Le passage à une acception continue de l'espace consiste aussi en un changement de mode d'implantation

¹²⁶ <http://www.hypergeo.eu/spip.php?article274> (Texte de G. Palsky, consulté le 22 octobre 2012).

spatiale : depuis le mode surfacique vers le mode ponctuel, par le biais d'une méthode d'interpolation des données qui permet de combler les lacunes en simulant les données.

Ce nouveau paradigme de la représentation continue de données qui sont discrètes, de part leur mode de collecte, autorise le transfert de méthodes issues de la mécanique des fluides à la cartographie statistique (Tobler, 1978 ; Tobler, 1981 ; Dorigo et Tobler, 1983). Sur le plan cartographique, l'avantage des procédures d'interpolation est d'éviter les problèmes liés au placement et à la disposition des figurés qui illustrent les flux, et ceux liés à leurs positions arbitraires au centre des unités spatiales (nous reviendrons sur ce point dans la section suivante, voir section 4.2.2). Sur le fond, l'approche continue permet de compenser d'éventuels effets de bords aux frontières externes de la zone d'étude liés à l'absence de données. Il convient de noter que cette acception continue de l'espace n'est pas le seul fait de W. Tobler : ce dernier présente, d'ailleurs, des initiatives similaires de ses contemporains des années 1970.

Le transfert de méthodes issues de la physique à l'analyse (cartographique) des flux est fructueux. Il constitue en premier lieu une bonne réponse à *l'effet spaghetti*, en proposant des solutions mathématiques à un problème graphique. L'approche continue est, en effet, bien adaptée à la cartographie des matrices denses en ce sens qu'elle produit un résumé de la structure spatiale des flux (Tobler, 1987 : 157) sans avoir à opérer de sélection *a priori* des éléments à représenter. En outre, parce qu'elle a permis à l'auteur de formaliser mathématiquement les lois empiriques sur les migrations proposées par E.G. Ravenstein (1885,1889).

4.1.5.4 Et la boucle est bouclée : de Tobler à Ravenstein

La contribution de W. Tobler à la cartographie de flux porte également sur la formalisation mathématique des sept lois de la migration de Ravenstein (1885-1889) au sein d'une théorie des migrations qu'il qualifie de *Push-pull migrations laws* (Dorigo, Tobler, 1983). L'idée de départ consiste à formuler de manière algébrique les lois empiriques de la migration proposées par l'auteur (voir Encart 4-1, page 37) puis à mettre en évidence la structure spatiale interne des matrices de flux OD issues de recensements de population.

W. Tobler considère, par exemple, qu'une caractéristique essentielle de la structure spatiale des flux migratoires est liée à la proximité des lieux : les flux sont d'autant plus importants entre deux zones que la distance qui les sépare est faible, ce qui traduit l'effet négatif de la distance sur l'intensité des échanges. La démonstration de ce trait particulier des phénomènes migratoires l'a notamment conduit à la formulation de sa première loi de la géographie. Sur le plan cartographique, ce sont ces recherches menées par l'auteur sur la modélisation des interactions spatiales, dans le cadre de la proposition d'une version personnelle du modèle d'interaction spatiale, qui ont conduit l'auteur à proposer la méthode de décomposition de la matrice de flux en deux composantes (voir chapitre 3).

Les travaux de W. Tobler ont conduit à des avancées majeures en matière de cartographie statistique des matrices de flux, tant sur le plan du traitement des données numériques que cartographiques, au préalable de leur représentation. Pour leur mise en œuvre, l'auteur a développé des outils qui ont par ailleurs contribué à l'émergence d'une nouvelle discipline : la géomatique.

Sans rentrer dans les détails de l'apport des travaux de W. Tobler au développement des sciences de l'information géographique, on observe depuis une bonne vingtaine d'année le développement intensif de l'informatique comme outil de traitement des données, simultanément de l'écran comme support des représentations. Si l'ère de l'Informatique graphique (*InfoViz*), que l'on pourrait qualifier de graphisme assisté par ordinateur, est surtout le fait d'informaticiens, les développements qu'ils ont réalisés constituent une source indéniable de renouvellement des aspects technologiques liés à la cartographie de flux.

4.1.6 A l'heure de l'Informatique graphique : vers le renouvellement de la sémiologie

La recherche théorique en cartographie de flux a désormais intégré les apports de l'informatisation, d'ailleurs largement initiés par les travaux¹²⁷ de W. Tobler puis l'engouement semble s'être tassé. J.C. Müller (1991) présente, dans un article déjà cité, dans l'introduction générale, sur *Le problème des visualisations des migrations inter-états aux Etats-Unis* une illustration des méthodes informatisées de cartographie de flux, soulignant ainsi la banalisation de solutions plutôt rudimentaires, souvent fondées sur l'analyse statistique des correspondances. Les représentations qui en résultent forment le « *modèle des atlas* » évoqué dans le chapitre 2. Nous rappelons que l'auteur prône le recours à la matrice ordonnée en complément de méthodes graphiques et mathématiques pour visualiser les flux. L'informatisation a, en effet, conduit à un renouvellement indéniable des méthodes de cartographie, avec le développement des Systèmes d'Information Géographiques (SIG), mais également à leur standardisation. Peu de modèles graphiques, de variables visuelles et de procédés graphiques sont implémentés dans les outils existants, les flux apparaissent comme le parent pauvre. On constate néanmoins, depuis quelques années, un regain d'intérêt pour la visualisation des données relationnelles, en particulier la prise en compte de l'espace dans l'analyse des réseaux à travers le concept de réseau spatial (*spatial network*)¹²⁸ (Barthelemy, M., 2010) qui est susceptible de profiter à l'analyse cartographique des flux.

Cet engouement émerge en lien avec trois processus simultanés. Le premier tient à l'explosion de masses de données (*big data*) libres, mises à la disposition du grand public, dans le cadre d'un processus général d'ouverture des données (*opendata*). Les bases de données qui décrivent des flux ou des réseaux de relations spatialisées sont, en effet, de plus en plus facile d'accès que ce soit à l'échelle locale, régionale ou globale. Les progrès du champ de la visualisation, notamment interactive, simultanément à l'ouverture des données, contribuent aussi à mieux prendre en compte le temps dans les nouvelles représentations cartographiques (Cheylan, 2007), à mieux représenter les mouvements. Un cas typique de cette tendance est celui des flux aériens internationaux ou des réseaux d'acteurs dont les représentations sont magnifiées. Les efforts sont, en effet, majoritairement réalisés sur les plans techniques et technologique des outils, tant pour la mise à disposition d'interfaces de visualisation que dans le développement de bibliothèques d'algorithmes, D3.js par exemple, permettant la réalisation aisée de cartes interactives. Une bonne présentation des dernières possibilités de géovisualisation est disponible dans un article récent de L. Jégou (2014), paru dans la revue *M@ppemonde*. Dans quelques cas, l'effort porte également sur le développement de méthodes de discrétisation qui prennent en compte le temps (Antoni, Klein, Moisy, 2012) ou sur la proposition de variables visuelles dynamiques (Andrienko, Andrienko, 2004 ; Kaddouri, 2008 ; Zanin, Ben Rebah, 2011) adaptées à une représentation interactive des flux, dans l'objectif d'augmenter la séparation visuelle des figurés pour augmenter la lisibilité des figures sur un support écran.

La pratique de la cartographie statistique contemporaine a donné lieu à différentes astuces graphiques qui tentent d'améliorer l'esthétique et le rendu visuel de la carte. L'utilisation de certains procédés ont été généralisés, parce que théoriquement plus faciles à mettre en œuvre, avec le support écran. C'est en

¹²⁷ L'auteur est le premier à avoir développé un logiciel de cartographie des flux pour ses propositions (Tobler, 1987). Le *Tobler's FlowMapper* est téléchargeable à l'adresse suivante (<http://www.csiss.org/clearinghouse/FlowMapper/>). Ce logiciel a ensuite fait l'objet d'un nouveau développement pour la gamme d'outils d'ArcGis, dans le cadre de la thèse de A. Glenon en 2006. Le fichier *.mxd pour ArcMap est téléchargeable à l'adresse : <http://dynamicgeography.ou.edu/flow/>. On regrettera que ce dernier ne soit pas compatible avec la série 10x de la gamme ArcGis. Dans le même ordre d'idées, on cite l'initiative peu aboutie et encore fermée au cas des flux chinois, de Danhuai, G. (2011) qui propose un outil interactif disponible à l'adresse : <http://cybergis.cn/cn/fmapp/fm.html> et fortement inspiré des solutions proposées par W. Tobler et D. Dorling, et de l'architecture de A. Glenon (2006).

¹²⁸ Ce concept fait référence à la prise en compte de l'influence de la position des nœuds ou des liens dans l'espace sur la structure et l'organisation du réseau ; il concerne d'abord l'analyse des flux matériels qui s'appuient sur une infrastructure (voies de chemin de fer, routes, conduites d'énergie notamment)

particulier le cas de deux problèmes graphiques : le premier tient à la gestion de la défonce graphique et le second au degré de généralisation des contours. Dans le premier cas, qui n'est pas spécifique à la cartographie de flux, la solution proposée par D. Dorling (1991)¹²⁹ consiste à détourner les flèches afin de renforcer l'effet visuel de séparation. Dans le second cas, la solution de W. Tobler (1987) consiste à introduire un effet 3D¹³⁰ ou à généraliser les contours, cette dernière ayant d'ailleurs été automatisée¹³¹ récemment.

Un certain nombre de développements ont été réalisés sur les aspects purement graphiques de la représentation des matrices. Deux directions sont particulièrement intéressantes pour la cartographie de flux : il s'agit des actions menées au niveau du tracé des figurés, en particulier des liens ; ainsi que celles qui permettent d'agir sur leur teinte, dans l'objectif de mettre en évidence les motifs des flux.

4.1.6.1 *Le tracé des liens : de l'arbre hiérarchique spatialisé à « l'edge bundling »*

Sur le plan graphique, la nouveauté de la série de propositions émanant d'informaticiens est fondée d'une part, sur l'utilisation de l'arbre hiérarchique pour représenter les matrices et d'autre part, sur l'utilisation de traits courbes (courbe de Bézier) à la place de traits droits.

Le fait d'incurver les liens est, en effet, apparu comme une solution efficace à la géovisualisation des masses de données relationnelles. Pour cela, les auteurs ont développé des méthodes d'interpolation des tracés directs (origine-destination) entre les lieux, dans la recherche d'une harmonie visuelle. Plusieurs versions de la méthode sont proposées dans la littérature (Holten, 2006, 2009 ; Lambert, Bourqui, Auber, 2009 ; Lambert, Aubert, Mélançon, 2010), dans l'objectif de réduire voire de supprimer l'effet spaghetti.

K. Verbeek et al. (2011) indiquent à ce sujet que la méthode de représentation des flux sous la forme d'« arbre en spirale logarithmique » qu'ils proposent comme variante de l'arbre hiérarchique, proposé par Phan et al. (2005), permet de supprimer « en douceur » l'effet de chevauchement des liens (donc, la partie de l'effet spaghetti liée à la surcharge graphique), sans opérer de traitement particulier préalable des données. Elle s'appuie sur une variante de l'arbre de Steiner¹³² ainsi défini par les auteurs : « *Un arbre (géométrique) Steiner pour un ensemble P de points du plan est une ligne courbe de longueur minimale qui relie tous les points de P soit directement, soit par l'intermédiaire d'autres points dits de Steiner. On peut montrer qu'un tel graphe est toujours un arbre traversant libre* ». Le résultat produit, effectivement, une image aérée des flux qui minimise les chevauchements entre les liens, mais également entre les liens et les différents éléments de la carte réduisant ainsi ce que nous appelons l'effet de couverture spatiale.

La Figure 4-15 présente quelques images caractéristiques de cette méthode (D. Holten 2006, 2009, Lambert et al., 2010). La figure classique des flux réalisés avec un trait droit correspond à l'image (a) ; les autres images (b, c et d), présentent des variantes de l'edge bundling où le trait est courbe.

¹²⁹ Une image de ces flux est consultable à l'adresse : <http://www.sasi.group.shef.ac.uk/thesis/chapter6.html>

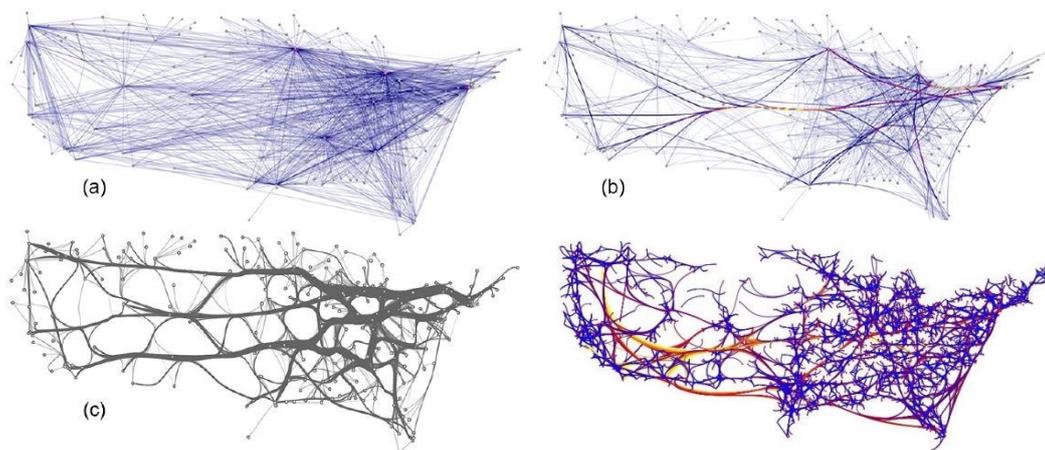
¹³⁰ Les astuces liées au degré de généralisation des contours des unités spatiales du fond de carte et du rendu 3D ont été proposées dans les années 1980. Tobler (1987, 159) a par exemple suggéré l'affectation d'une teinte grisée sur les versants sud-est de la carte afin de donner l'illusion d'un effet 3D. Ce procédé de 3D est également, utilisé, notamment pour déformer les unités spatiales par rapport à une quantité : c'est le cas des cartes dites « prismatiques ». Concernant la représentation de données qui portent sur des mobilités spatiales, le résultat n'est toujours très heureux. L'effet 3D ne contribue pas à accroître la lisibilité de la figure : la carte réalisée par Ratti et al. (2010) est ce sens symptomatique d'un usage abusif de l'effet-3D, tout comme de la couleur.

¹³¹ Voir à ce sujet, l'application Mapshaper (<http://www.mapshaper.org/> et <https://github.com/mbloch/mapshaper>) développée par M. Bloch sous la supervision notable de M. Harrower, qui permet de généraliser aisément un fond de carte.

¹³² L'arbre de Steiner, du nom d'un mathématicien du même nom, s'apparente à la notion d'arbre couvrant minimum qui vise à tracer un graphe formé des seuls liens qui relient l'ensemble des sommets. L'algorithme de Kruskal, souvent utilisé pour visualiser les réseaux sociaux est une méthode d'extraction d'arbre couvrant minimum.

On observe ainsi que la méthode ne modifie la figure que dans son apparence visuelle, dans son esthétique et non dans son fondement. Pour extraordinaire que ces figures puissent apparaître, il est intéressant de noter qu'elles s'inspirent de cartographies anciennes de routes voir (Figure 4-16) – non de flux – développées au début du XX^{ème} siècle et réalisées pour l'essentiel à la main.

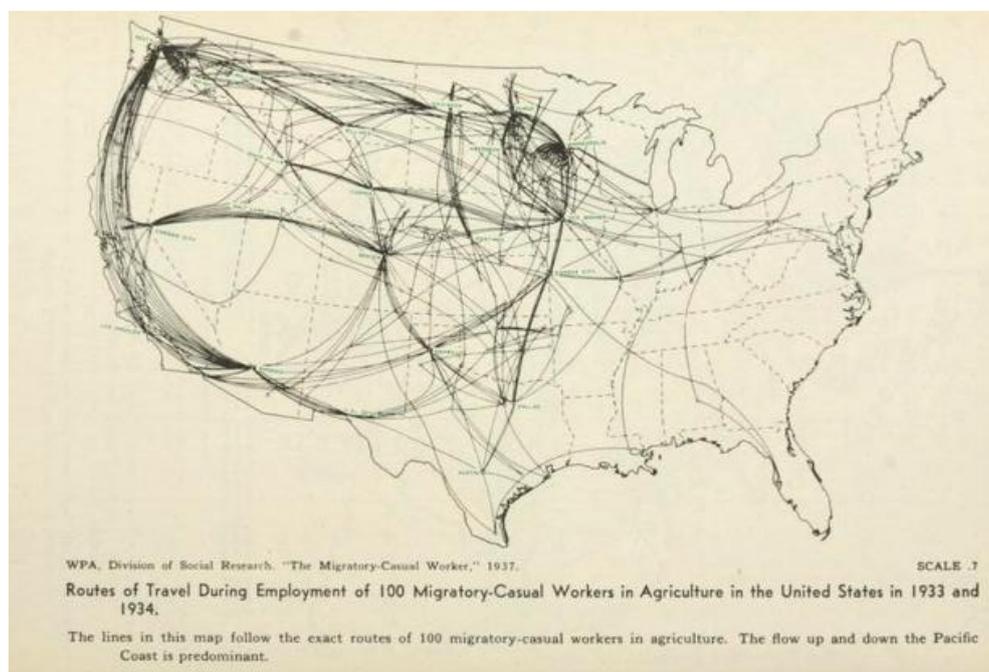
Figure 4-15. Quelques figures actuelles de la carte de flux (~2010)



Sources : (a), (b) et (c), D. Holten, J. Van Wijk (2009) ; Lambert, A., et al. (2010)

Ces figures ont été réalisées à partir de matrices de type (L_{ij}) pour l'image a) et de type (F_{ij}) pour les autres images. L'objectif des auteurs étant focalisé sur la visualisation d'une densité importante de liens [image (a)] éventuellement valués et la perception de motifs qui correspondent à l'orientation générale des flux, à la mise en évidence de routes préférentielles.

Figure 4-16. Une figure de la carte de flux (1934)



Source: Brinton, W-P.,(eds.), 1939.

Cette carte décrit les déplacements de travailleurs agricoles saisonniers des années 1930, aux Etats-Unis, en suivant leur tracé exact. La légende mentionne, en effet, que les lignes (L_{ij})

suivent les routes exactes de ces migrants saisonniers agricoles ; les flux montant et descendant de la côte pacifique sont prédominants. En fait, les lignes courbes correspondent, à la représentation l'orientation générale des routes empruntées par ces travailleurs plus que leurs routes réelles.

La Figure 4-16, réalisée en 1934 pour illustrer les flux de travailleurs migrants dans l'agriculture américaine est caractéristique de ces cartes anciennes décrivant des routes. Elle illustre parfaitement à notre sens l'ancienneté des méthodes de cartographie de flux qui font aujourd'hui l'objet d'un renouvellement visuel rendu possible par l'informatique graphique, elle montre l'intérêt des outils informatiques actuels dans l'actualisation de méthodes anciennes et dans la facilitation de leur mobilisation.

Sur les images de la Figure 4-15, les liens qui illustrent les flux sont fusionnés graphiquement – et donc visuellement – de manière à ce qu'ils apparaissent comme suivants des routes préférentielles telles qu'elles étaient représentées il y a près d'un siècle. On observe aussi que ces deux figures (Figure 4-15 et Figure 4-16) jouent de la superposition des liens et de leur proximité dans le plan, afin de renforcer la perception des motifs principaux, en particulier les directions ou bien les zones où il se passe plus de choses qu'ailleurs. Ces zones sont présentées dans des teintes qui apparaissent plus foncées, grâce à l'astuce de superposition. On devine également un jeu de lumière qui conduit l'observateur à deviner, par l'effet de transparence induit – et de superposition – les liens ou les lieux qui ont une position moindre dans le système global.

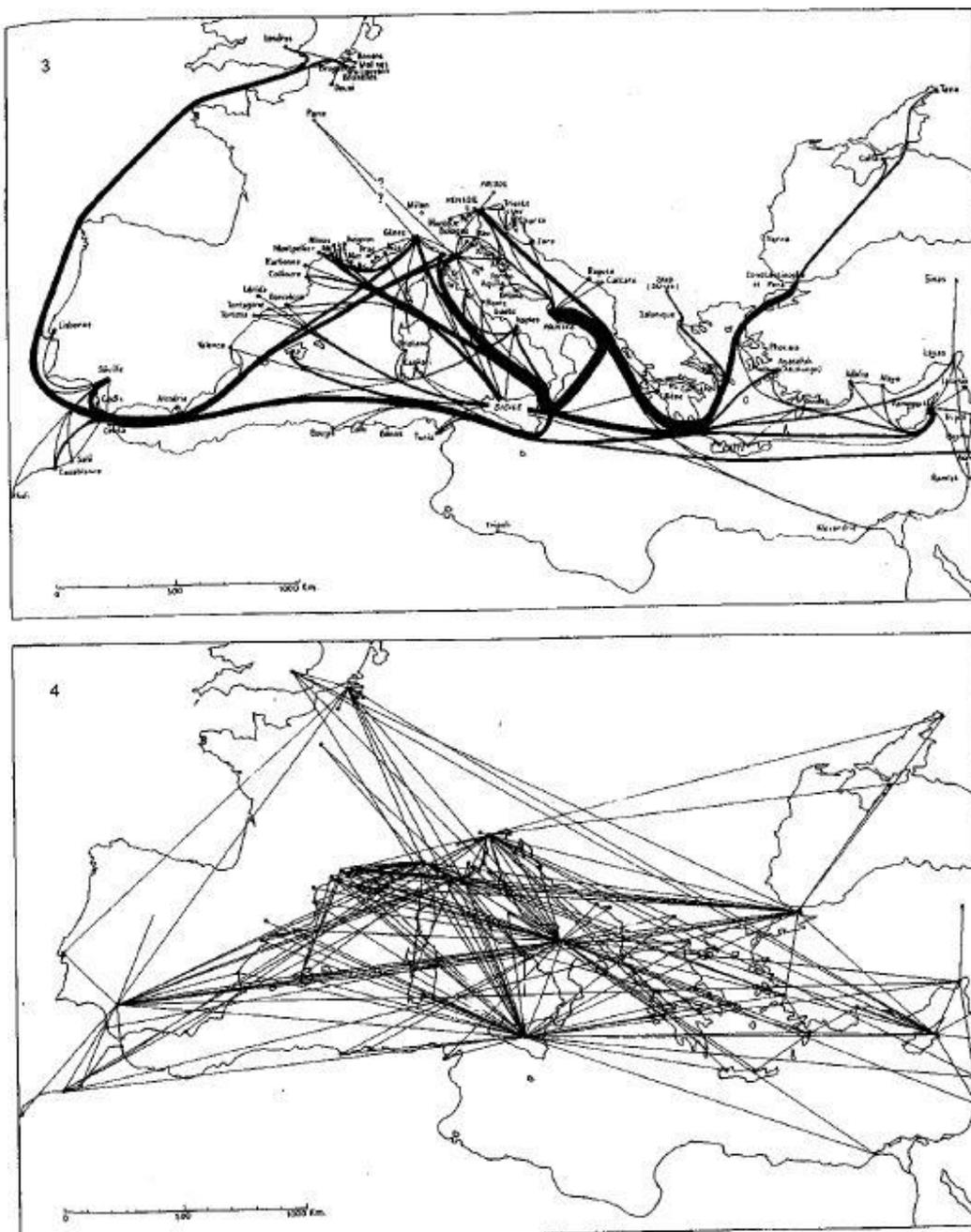
Il est par conséquent très intéressant de noter la ressemblance frappante entre ces deux cartes (Figure 4-15 et Figure 4-16). Les premières réussissent donc la double prouesse de remettre au goût du jour des méthodes anciennes de représentation des densités de chemins ; la différence entre ces deux figures tient, peut-être et encore nous n'en sommes pas sûre, dans la densité de liens qui semble moindre sur la figure Figure 4-16.

Si ces méthodes de fusion graphique renouvèlent, indéniablement, l'image actuelle des relations spatialisées, il convient de noter que le modèle graphique auquel elles semblent appartenir n'est pas nouveau. En témoigne cette proposition de J. Bertin de représenter des relations non orientées que nous proposons sur la Figure 4-17.

La Figure 4-17 présente deux cartes de flux commerciaux de blé réalisés dans la méditerranée du XV^{ième} siècle (Bertin, 1973 : 345). L'auteur indique ce qui suit à leur propos. « (3) *Il ne suffit pas de tracer les itinéraires réellement parcourus pour représenter un système de relation. Une carte des routes maritimes, même pondérée, ne montre pas l'orientation commerciale des centres d'activité. Elle montre la densité des bateaux en mer. Les relations maritimes marchandes entre les villes de l'Europe et de la Méditerranée n'apparaissent seulement dans leur diversité, leur poids et leur orientation géographique que lorsque chaque relation, bien que maritime, est représentée par une droite* (4) [sur la carte du bas] » (Bertin, 1973 : 344).

Cette carte va se révéler très importante dans notre recherche pour deux raisons fondamentales : d'abord, parce qu'elle justifie l'idée de décomposer la matrice, ensuite parce qu'elle introduit la possibilité cartographique de représenter les flux de manière sinueuse et non plus rectilinéaire.

Figure 4-17. La cartographie de routes, seconde moitié du XX^{ème} siècle (Bertin, 1973)



Source : J. Bertin, 1973.

Ces cartes décrivent la même information : des flux commerciaux de blé. Sur celle du haut, la matrice correspondrait à $(Fij)^+$ tandis que sur celle du bas, à (Lij) représentant les relations entre ports méditerranéens. Dans le premier cas, les bandes sinueuses semblent suivre le tracé des routes maritimes réelles à l'inverse du second cas. Nous y reviendrons dans le chapitre 6.

La différence entre ces deux cartes de J. Bertin et celles de la Figure 4-15 tient à un élément particulier : une action sur la luminosité des figurés, d'ailleurs suggérée par (Bertin, 1973) pour gérer les superpositions, est rendue possible par les outils actuels du graphisme. En effet, l'action sur l'opacité et la luminosité de la teinte est devenu un axe majeur de développement graphique, dont nous

montrons dans le chapitre 5 la manière dont il contribue au renouvellement de l'image des flux. Il concerne principalement le volet esthétique de la figure lié à l'usage de la couleur.

4.1.6.2 L'usage de la couleur (teinte)

L'observation de cartes récentes réalisées dans le champ de la visualisation des réseaux spatialisés, nous a conduites à poser une hypothèse forte : la couleur, telle qu'elle est actuellement mise en œuvre dans le champ de l'informatique graphique, est susceptible d'améliorer la cartographie de flux dans ses aspects graphiques, liés à l'esthétique de la couleur en particulier. Nous montrerons dans le Chapitre 5 comment une utilisation de la teinte est susceptible de renouveler la sémiologie graphique de la carte de flux. La modification apportée s'inscrit non seulement dans le cadre d'un bouleversement esthétique, mais aussi d'un renouvellement de la mise en œuvre de la variable visuelle *Couleur*, conduisant, de notre point de vue, à renforcer la perception de certains motifs.

Le renouvellement de l'utilisation de la *Couleur* est rendu possible grâce à la généralisation du support écran qui a entraîné de nouvelles possibilités d'action sur ses propriétés de Teinte, de Saturation et de Luminosité (TSL), ayant conduit à la fameuse carte du réseau de relations des abonnés à Facebook© (Butler, 2010), présentée sur la Figure 4-18.

Figure 4-18. La carte des amitiés de Facebook©



Source : P. Butler, 2010.

La matrice qui se trouve à l'origine de cette carte est celle du réseau de relation (Lij) des abonnés au service de mise en relation Facebook. Elle renseigne sur l'existence de liens entre les villes d'enregistrement des abonnés, et celles de leurs amis. L'intérêt que nous portons à cette figure tient dans l'usage de la teinte des liens qui n'est pas uniforme le long du tracé. Elle est rendue lumineuse lorsqu'ils franchissent un espace continental et que la densité est importante ; à l'inverse, la teinte est sombre voire éteinte, lorsque l'espace est maritime, de sorte que l'on devienne le squelette du réseau de villes connectées au service. Outre les aspects liés à l'esthétique de la figure, qui sont indéniables, la question qui se pose est celle des choix de sémiologie cartographique liés à cet usage de la teinte, en l'occurrence de l'application du principe de l'opposition colorée à une telle matrice, dont on rappelle qu'elle n'est pas orientée. Nous y reviendrons en seconde partie, dans le chapitre 5.

La Figure 4-18 est typique de cette nouvelle esthétique de la carte de relations – non pas de flux dans le cas précis – réalisée et visualisée sur un support écran. Elle caractérise le fait que l'utilisation actuelle de la couleur ne se limite pas au simple transfert depuis le papier vers l'écran d'une

méthodologie ancienne. Il s'agit bien d'une nouvelle utilisation de la même variable vers une extension de ses possibilités.

La mise en œuvre de la couleur peut aujourd'hui être réalisée soit au niveau de la lecture d'ensemble de la figure pour caractériser un style, soit au niveau local de l'apparence des figurés et cela, dans un souci d'harmonie et d'esthétique visuelle qui est rendu possible grâce à deux applications disponibles librement en ligne : le *Color Brewer*¹³³ (Brewer, 2002) et le *geotest*, un outil d'exploration et de gestion des couleurs en cours de développement par L. Jégou. Cette mise en œuvre de la couleur porte sur le choix de la palette (et donc du style), sur la mise en œuvre du gradient coloré qui caractérise les nuances de teinte ainsi que sur l'utilisation de différentes propriétés, en particulier la transparence ou l'opacité, telle que suggérée par A. Mac Eachren (1995).

Nous examinons ci-dessous les deux dernières possibilités suggérées par A. Mac Eachren (1995) et mises en œuvre sur la Figure 4-18. La première consiste à utiliser un gradient coloré (noir – bleu – blanc) pour représenter les liens, en fonction d'une variable. Sur cette figure, la variable semble être à première vue, le franchissement d'une partie terrestre : les parties des liens concernées sont ainsi rendues plus lumineuses que celles qui franchissent les mers. Cette méthode semble être utilisée conjointement à une variation de teinte en fonction de la longueur du figuré : les liens terrestres les plus courts nous apparaissent comme les plus brillants, les plus lumineux par rapport à ceux qui sont plus longs. Cette luminosité apparaît en « opposition nette » avec les parties maritimes des mêmes liens qui sont, elles, très sombres, comme plongées dans la nuit. Sur le plan sémiologique, le résultat obtenu par la représentation de ces liens est saisissant : en rehaussant ainsi la perception visuelle des seules parties de liens qui sont terrestres, la carte en arrive à illustrer un réseau de lieux connectés. Elle renouvelle aussi l'esthétique de la carte de flux en différents points, parmi lesquels le style bleuté, la luminosité et courbure des liens. La méthode conduit aussi à la suppression de la flèche, en même temps qu'elle demeure particulièrement efficace pour transmettre l'information qui porte sur la mise en relation de deux lieux et la localisation des lieux les plus connectés. L'auteur a probablement souhaité mettre en évidence ce réseau local de villes connectées par le biais des abonnés au réseau social, ou montrer l'extension de l'accès à l'internet mondial de ceux qui sont ainsi « éclairés ».

La seconde possibilité d'utilisation de la couleur mise en œuvre sur la Figure 4-18 est liée à l'effet de transparence ou d'opacité dans le rendu des deux teintes utilisées. Une manière intéressante d'améliorer sensiblement la visualisation graphique des figurés sur un support écran consiste à agir sur le degré d'opacité de la teinte définie en (R,V, B) par exemple, en lui associant le canal alpha (R,V, B, A). Le canal alpha conduit, en effet, à une égalisation des critères de la teinte qui modifie l'apparence visuelle des objets sur lesquels elle s'applique.

Son utilisation est très intéressante lorsque la figure est rendue complexe, par des objets dont les valeurs – les nuances de teintes correspondantes – sont combinées (elles correspondent à deux séries de valeurs, par exemple) ou bien parce qu'ils sont superposés, ou qu'ils se chevauchent comme dans le cas de flux. La méthode consiste donc à ajouter une nouvelle propriété de nature morphologique à la *Couleur* qui facilite la perception de ces motifs complexes.

Elle conduit à la réalisation de cartes typiques, connues sous le terme anglais de *Value By Alpha (VBA) map*, une expression introduite d'après W. Tobler (n.d.)¹³⁴ par E. Raisz en 1938, réintroduite par A. Mac Eachren en 1995 pour désigner l'action de la variable visuelle *Transparence*. Il s'agit

¹³³ <http://colorbrewer2.org/>

¹³⁴ W. Tobler (n.d.) considère que le terme de "*Value By Alpha (VBA) map* est utilisé par Raisz (1938). Les soviétiques la qualifient de « varivalent projections », les français "d'anamorphoses [...] ».

d'une méthode informatique de *design* de la teinte formalisée en 2010 par R. E. Roth et al. (2010), en analogie avec l'utilisation de la *Taille*, lorsqu'elle est mise en œuvre dans le cadre d'une anamorphose, ce qui explique la proximité des dénominations. La VBA s'apparente à un clin d'œil à l'expression *Value By Area* (VBA) *map*, spécifique aux procédés d'anamorphoses scalaires qui déforment la morphologie des unités spatiales proportionnellement à la valeur d'un indicateur statistique. Ici, c'est plutôt la « morphologie » de l'arrangement des canaux (R, G, B) qui définit la teinte perçue qui est « déformée », par rapport à la valeur d'un indicateur. Le procédé conduit donc à un ré-échantillonnage des niveaux de teinte ce qui produit, au final, un double échantillonnage des mêmes données : au niveau de la valeur et au niveau de la teinte, pour un effet visuel saisissant.

L'intérêt de l'utilisation du canal alpha réside dans son applicabilité à chacune des unités spatiales qui caractérisent la zone d'étude, de la même façon que la *Taille* pour l'anamorphose. Une grande valeur alpha signifie que la teinte de l'objet est opaque, ou faiblement transparente et inversement. D'après les auteurs, le canal alpha de la teinte résout le dilemme de l'anamorphose (qui déforme la topologie et la morphologie des unités spatiales tout en autorisant une perception plus juste des motifs spatiaux), en égalisant parfaitement les teintes en fonction de l'objectif de la carte. Il est notamment possible d'agir sur la figure, en illuminant les unités spatiales qui présentent les valeurs les plus significatives (ou inversement), ce que les auteurs qualifient « *d'effet de projecteur* », ou d'ajuster la *Transparence* de manière à mettre en évidence un *effet de profondeur*, dont nous verrons qu'il est complémentaire à l'*effet d'amalgame* précédemment défini (voir section 3.3.1, consacrée à la cartographie des matrices asymétriques, page 128).

L'application de la *Transparence* conduit, en effet, à la déformation de la teinte, de sa perception de manière à faire apparaître comme des gradients qui correspondent aux besoins de représentation de données qui sont incertaines (Mac Eachren, 1992) ou qui sont superposées. D'abord mise en œuvre pour rehausser les cartes choroplèthes, la méthode est ensuite utilisée sur des cartes de visualisation des réseaux sociaux telle que celle de la Figure 4-18, caractérisées par un chevauchement important de figurés linéaires. C'est pourquoi nous posons l'hypothèse que **cette utilisation de la *Transparence*, en lien avec le choix de la teinte, est susceptible d'améliorer le rendu des cartes de flux, tant sur le plan esthétique qu'au niveau de la perception des motifs.** Nous examinerons cette possibilité dans le chapitre 5.

L'utilisation de la transparence des figurés ne suffira toutefois pas, à elle seule, à rendre toutes les cartes de flux lisibles. Une autre astuce a été utilisée : la gestion de la disposition des figurés. Sur la carte de la Figure 4-18, les liens les plus courts sont placés au premier plan, les différents figurés sont ainsi disposés de manière ordonnée sur la figure, en fonction de leur longueur. Cet ordonnancement des figurés fait écho à la suggestion de W. Tobler (1987) de disposer les figurés qui présentent les valeurs les plus importantes au premier plan de la figure.

C'est donc l'action combinée de la disposition des liens, d'un gradient de teintes et/ou de la *Transparence* qui permet d'augmenter la séparation visuelle des figurés qui illustrent les flux. Cette hypothèse diffère de celle de D. Dorling (1991), de détourner¹³⁵ les liens, c'est-à-dire d'appliquer le procédé de la défonce graphique aux flux, en leur affectant un liseré blanc. A noter ici que cette proposition signifie qu'il faille illustrer les flux par des bandes, dont la surface est mesurable, et non

¹³⁵ La suggestion de D. Dorling (1991) de détourner les flèches, pour mieux les différencier, a été mise en œuvre dans le cadre d'une cartographie interactive plutôt que dynamique. Elle est couramment mise en œuvre dans les cartes choroplèthes qui représentent les variations des distributions des lieux par des cercles proportionnels. Ici, les petits symboles sont d'autant plus visibles qu'ils sont renforcés par un contour blanc et placés au dessus des plus gros. Il s'agit d'une variante du détourage (Tobler, 1987 : 159) qui consiste à augmenter le degré de séparation entre le symbole et le fond de carte, en créant un espace.

par un trait ou une courbe : le liseré ne pouvant être appliqué qu'à une forme surfacique qui présente un fond et un contour.

Il convient de noter que ces deux possibilités mises en œuvre sur la Figure 4-18, relèvent surtout du *design* graphique de la couleur, c'est-à-dire de la teinte. Elles ne concernent pas la mise en œuvre de la variable visuelle *Couleur* sur laquelle nous reviendrons dans le chapitre 4, en lien avec ces nouvelles propositions.

A l'heure de l'informatique graphique, les avancées dans le domaine de la visualisation des réseaux foisonnent. Elles sont surtout le fait de développeurs (majoritairement informaticiens) plutôt que de thématiciens (des géographes). Les exemples que nous avons proposés n'ont pas été réalisés, à notre connaissance, en collaboration avec leurs potentiels utilisateurs. Leur efficacité et surtout leur utilité est, par conséquent, amoindrie puisque les outils qui permettent leur mise en œuvre sont soit difficiles à prendre en main lorsqu'ils sont accessibles, soit inadaptés aux analyses réalisées en sciences humaines et sociales, lorsqu'ils ne génèrent pas d'erreurs, soit focalisés sur les seuls aspects liés au *design* au détriment des aspects méthodologiques et cartographiques.

C'est pour cela que T. Joliveau¹³⁶ (2011) et M. Baron et al. (2012) ont proposé des analyses critiques de ces nouvelles images des réseaux (parmi lesquelles la Figure 4-18) qui sont le témoin, de « [...] *l'appropriation de l'outil cartographique par des acteurs qui n'ont pas fait l'effort de s'y former méthodologiquement et théoriquement. Car la représentation juste d'une telle quantité de données sur un si grand espace n'est pas chose aisée. C'est même un tour de force cartographique. [...] La seconde remarque concerne l'objectif principal de ces cartes, qui se limite à la promotion de la valeur esthétique du résultat aux dépens de l'efficacité cartographique* » (Baron et al., 2012).

Ces figures sont souvent le résultat de l'application d'algorithmes de dessin, d'outils développés d'abord dans un objectif esthétique, pas théorique ou méthodologique. Ce constat n'est pas spécifique à ces images : les outils de la cartographie statistique, lorsqu'ils sont apparemment appropriés par les développeurs de solutions informatiques, ne répondent pas particulièrement aux besoins des cartographes. Pour preuve, l'application de géovisualisation des flux *Jflowmap*© (Boyandin et al., 2010), pourtant la plus aboutie de notre point de vue, présente différents problèmes¹³⁷ d'ordre méthodologique qui ne facilitent pas sa prise en main. Pour autant, l'outil présente d'autres qualités indéniables à commencer par la mise à disposition libre du code.

Doit-on voir dans ces quelques lignes la résurgence d'un reproche récurrent fait aux enseignements universitaires en (méthodes de) cartographie qui ne conduisent pas, pour l'essentiel, au développement d'une recherche théorique sur les méthodes de représentation ? Et pourtant, outre les questions liées au traitement des données, il n'est sans doute pas inutile de rappeler d'une part, les principes de sémiologie graphique qui s'appliquent aux flux, tels qu'ils ont été rassemblés et formalisés par J. Bertin depuis 1967. Et d'autre part, les aspects de la conception cartographique liés au tracé des figurés qui illustrent les flux et qui se trouve à l'origine de nombreuses contraintes ayant contribué à l'évitement de la carte de flux.

4.2 La conception d'une carte de flux : aspects graphiques et sémiologiques

Puisqu'il n'est pas envisageable de tout dire sur tout et, en matière de sémiologie graphique, il y en aurait beaucoup, cette section ne portera pas sur l'évaluation des méthodes et des procédés de visualisation, ni sur les aspects sémiotiques ou cognitifs de perception des formes. Cette section porte

¹³⁶ <http://mondegeonumerique.wordpress.com/2011/01/10/500-millions-damis-la-carte-de-facebook-1-deconstruction/>

¹³⁷ On peut citer: une gestion difficile voire impossible des systèmes de projection cartographiques les plus courants, la génération de points qui ne sont pas proportionnels aux valeurs représentées et qui correspondent, nécessairement, aux données disponibles dans la partie diagonale de la matrice ; l'absence de compatibilité avec les autres outils nécessaires à la réalisation de graphiques, lors de l'export

plus largement sur la conception d'une carte de flux : elle a pour objectif de présenter les figurés qui existent ainsi que leur agencement sur une planche cartographique, en tenant compte des principes de sémiologie graphique qui conditionnent leur mise en œuvre, que nous présentons dans la première sous-section. Cette section propose ainsi un état de l'art des méthodes de conception des cartes de flux qui tient compte des apports des différents auteurs présentés dans la première section de ce troisième chapitre.

A la suite des cartes de H. D. Harness (1836) et de E. G. Ravenstein (1885), le début du XX^{ème} siècle est marqué par une pléthore de procédés de cartographie de flux, tous aussi uniques les uns que les autres, en témoigne le nombre de synthèses disponibles. La plus ancienne, à notre connaissance, est l'Atlas *Graphic Presentation*¹³⁸ des modes de représentation et procédés graphiques publié par l'éditeur Brinton en 1939, en l'honneur de W. Playfair considéré comme l'inventeur de la statistique graphique. En 1962, W. Bunge (1962, 1966) propose également dans son ouvrage *Theoretical Geography*, une synthèse des différents procédés de représentation du mouvement.

Trente plus tard, D. Dorling publie en ligne¹³⁹ en 1998 une bibliographie sur la visualisation de données dans les sciences sociales, issue de sa thèse de doctorat, *The Visualisation of spatial social Structures*, qu'il soutient en 1991 à l'Université de Newcastle. Cette thèse sera finalement éditée en 2012, sous une forme papier (Dorling, 2012) et électronique¹⁴⁰.

On observe plus récemment deux initiatives remarquables de compilation des modèles et procédés. La première émane du groupe de travail dirigé par la documentaliste K. Börner (2011) qui a notamment conduit à la réalisation d'un atlas sur la visualisation de nos connaissances (Bahoken, 2012) ; elle est complétée d'un site internet¹⁴¹ et par des expositions itinérantes intitulées *Places et Spaces: Mapping Sciences*. La seconde initiative notable est celle de l'historienne S. Schulten (2012) qui passe en revue, dans l'ouvrage *Mapping the Nation*¹⁴², la progression des méthodes de représentation cartographique dans la culture américaine. Ces deux initiatives récentes de compilation des procédés de représentation cartographique sont la marque d'un regain d'intérêt général pour les techniques et méthodes de la représentation graphique des données quantitatives. On notera toutefois que les références les plus récentes ne portent pas sur les aspects purement cartographiques, ceux qui sont liés à la carte d'une matrice, ce qui confirme notre première hypothèse selon laquelle la carte de flux n'est qu'un cas particulier¹⁴³ de graphe. Historiquement et sur le plan formel (Palsky, 2004 b), les graphes ont d'ailleurs précédé les cartes de flux, ils leur ont ouvert la voie.

L'efficacité d'une carte de flux réside dans la retranscription et la suggestion de la dynamique des mouvements, ce qui correspond à un véritable défi. Du point de vue de la conception de la figure, celui-ci suppose de prendre en compte « - le déplacement : c'est-à-dire le changement de localisation d'un mobile ; - le facteur temps¹⁴⁴ : vitesse rapide ou lente, constante ou variable ; - le volume ou la quantité d'énergie déplacée : flux, force, débit » (Weger, 1999 : 122).

¹³⁸ Une version allégée du document complet, peut être téléchargée, au format PDF (30Mo), à l'adresse suivante :

<http://www.archive.org/details/graphicpresentation00brinrich> [lien vérifié le 02/08/2015]. Voir notamment les pages 216 à 230 pour les cartes de flux, et 73 à 80 pour les diagrammes de flux.

¹³⁹ <http://www.agocg.ac.uk/train/review/exec.htm> [lien vérifié le 02/08/2015].

¹⁴⁰ <http://www.dannydorling.org/books/visualisation/Graphics/Pages/Figures.html> [lien vérifié le 02/08/2015].

¹⁴¹ <http://scimaps.org/> [lien vérifié le 02/08/2015].

¹⁴² Cet ouvrage est présenté en ligne, à l'adresse suivante <http://www.mappingthenation.com/> [vérifiée le 02/08/2015].

¹⁴³ La carte de flux n'est pas considérée comme un modèle graphique dans la classification sous la forme de table périodique proposée par le site internet *Visual Literacy*. Celle-ci est disponible en ligne, à l'adresse suivante : http://www.visual-literacy.org/periodic_table/periodic_table.html [lien vérifié le 02/08/2015].

¹⁴⁴ Le temps dont il est question doit être compris comme d'une part, le niveau des données (intégration du facteur vitesse, donc du temps du phénomène) ; de l'autre, celui de la représentation (animations sur support écran).

Un second défi consiste à dépasser les limites inhérentes au format statique pour figurer le mouvement et son orientation, selon les principes de la graphique de communication appliqués à la cartographie de flux. Ces principes sont mis en œuvre compte tenu des contraintes graphiques inhérentes à la cartographie de flux.

4.2.1 Aspects sémiologiques de la cartographie de flux

Les aspects sémiologiques de la cartographie de flux intéressent la graphique de communication. Il s'agit d'une branche de la visualisation graphique de données statistiques qui utilise un alphabet graphique, les variables rétinienne ou visuelles, pour communiquer un message. La graphique de communication diffère donc de la graphique de traitement qui concerne les traitements matriciels tels que, par exemple, le ré ordonnancement des lignes et des colonnes de la matrice pour aboutir à une matrice ordonnée telle que nous l'avons présentée dans la section 1 du Chapitre 2. Ces variables visuelles sont présentées dans la première sous-section, elles sont appliqués sur des figurés dont l'apparence ou la forme fait l'objet de la seconde sous-section.

4.2.1.1 Les variables visuelles mobilisées

Les variables visuelles appartiennent à un projet de sémiologie graphique développée pour l'essentiel au XIX^{ème} siècle. F. Saint-Martin (1997) attribue à F. de Saussure, linguiste suisse du début du XX^{ème} siècle, l'idée de développer cette science générale fondée sur l'étude des systèmes de signes et de leur signification, une sémiologie où le langage oral ne serait qu'un élément de l'ensemble.

L'alphabet (le point, la ligne et la tâche) sont agencés pour former des mots (les figurés) dont la composition selon certains principes forme les variables visuelles. Ces principes sont répertoriés et formalisés dans un ouvrage de référence publié en 1967 par J. Bertin (1967, 2005). L'auteur a identifié six variables visuelles : la *Taille*, la *Valeur*, le *Grain* (et la *Texture-Structure*)¹⁴⁵, l'*Orientation*, la *Forme* et la *Couleur* qui sont toujours considérées comme fondamentales. La mise en œuvre de ces variables varie en fonction de l'implantation graphique de l'objet et donc de l'unité spatiale considérée selon qu'elle est ponctuelle, linéaire ou surfacique.

La cartographie de flux, en mobilisant des objets de différentes implantations est par conséquent concernée par l'action conjointe de plusieurs variables visuelles. Cependant, toutes ne revêtent pas un caractère fondamental pour la représentation et la compréhension des flux représentés. Les variables visuelles fondamentales sont, en fonction du type de matrice : la *Taille* et la *Couleur* appréhendée comme une nuance de teinte (perçue en fonction de la luminosité) et comme une variable d'échantillonnage, en association avec la *Valeur*, pour traduire la proportion de la teinte sur une palette. Il s'agit généralement d'une part de la part de teinte noire du figuré tracé sur un fond blanc.

Le corpus initial des variables visuelles a fait l'objet, ces dernières années d'extensions par différents auteurs. Ces extensions peuvent être considérées, soit comme une nouvelle combinaison des variables existantes, soit comme une nouvelle façon de les mettre en œuvre individuellement. Elles ont été majoritairement impulsées par les nouvelles possibilités de visualisation offertes par le support écran, par les outils de Cartographie Assistée par Ordinateur (CAO) puis des Systèmes d'Information Géographique (SIG) devenus aujourd'hui incontournables pour leur mise en œuvre aisée (Skupin, 2000).

¹⁴⁵ D'après C. Zanin, la variable visuelle *Grain* est assimilée à la *Texture*, à l'origine des principes de sémiologie graphique. Les deux termes sont cependant distingués, dans la pratique, par J. Bertin : le *Grain* serait lié aux changements de texture de l'image (du grain), à la manière d'une focale photographique ; la *Texture*, appliquée en association à la *Forme*, permettrait de remplacer l'idée de trame.

Parmi les directions notables dans la littérature scientifique, nous pouvons notamment mentionner trois familles de propositions d'extensions des variables visuelles traditionnelles qui sont susceptibles d'intéresser la cartographie¹⁴⁶ statistique de flux.

La première porte sur l'intégration du temps ou de la dynamique des phénomènes (Mac Eachren, 1992, 1994, 1995 ; Kaddouri, 2008 ; Zanin, Ben Rebah, 2011) dans la représentation, principalement en lien avec les possibilités d'animation de la figure et d'interactivité du support. La seconde famille concerne la modification du tracé du figuré linéaire¹⁴⁷ qui illustre le flux : le choix de sa texture ou de son *Grain*, la manière de représenter son *Orientation*, sa *Taille* et, plus généralement l'arrangement de l'ensemble à l'échelle globale de la figure, de leur densité (Caivano, 1990 ; Dorling, 1991 ; Tobler, 1987 ; Dent, 1999). La troisième famille est liée à la mise en œuvre de la *Couleur*. Une première direction porte sur la combinaison de nuances de teintes, en jouant ou non sur leur degré de saturation (Caivano, 1990). Une seconde direction est impulsée par les travaux d'A. Mac Eachren (1995) qui suggèrent d'ailleurs, à la suite de C.-J. Minard (voir section 4.1.2.2, page 148), une action sur l'opacité de la teinte des figurés qui correspond à leur degré de *Transparence*, introduisant alors une variable visuelle éponyme. L'auteur a, en effet, bien montré la manière dont le niveau de luminosité de la teinte du figuré pouvait influencer sur la manière dont il est perçu, sur sa netteté (Mac Eachren, 1992, 1995).

Ces variables visuelles sont appliquées à la représentation de flux, plus précisément au figuré qui représente le flux, dans le cadre d'une adaptation rendue nécessaire par les spécificités de la représentation d'une information relationnelle.

4.2.1.2 Le tracé du figuré qui illustre le flux : le trait

Le figuré qui illustre le mieux le flux est un trait dessiné en implantation linéaire, éventuellement orienté. A. Mac Eachren (1979) précise que le trait a d'abord été mis en œuvre pour représenter la direction d'un mouvement, avant d'être rendu plus complexe en intégrant le volume ou la quantité en mouvement, par une variation de ses dimensions (principalement sa *largeur*) et de sa direction.

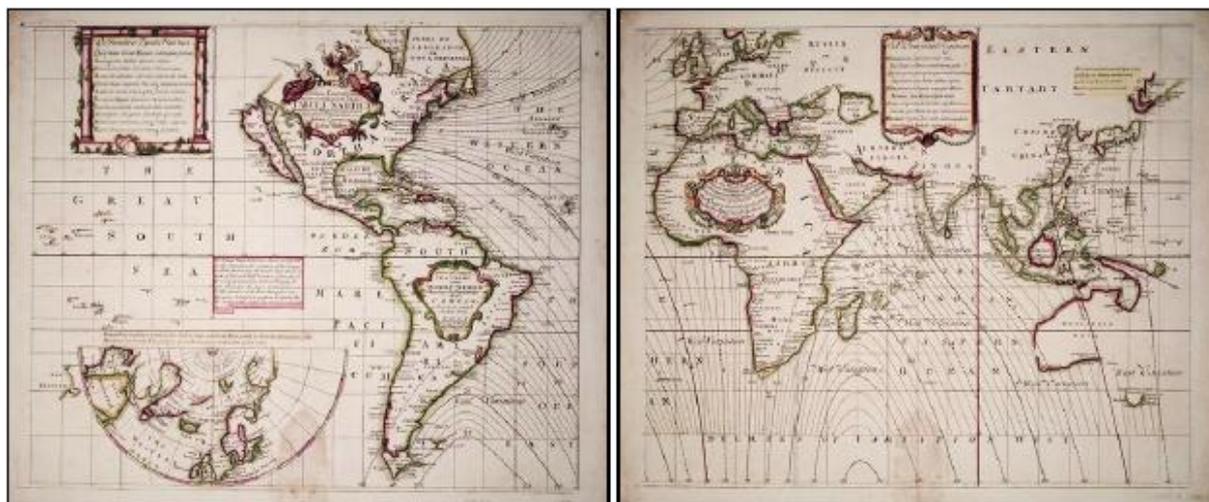
Le symbole linéaire, représenté par la flèche est, en effet, celui qui convient le mieux pour représenter des flux (Bertin, 1967, Béguin, Pumain, 1994 ; Dent, 1999, Poitevin, 1999, Zanin, Trémélo, 2003). Il est utilisé pour représenter tant des flux quantitatifs mesurables que qualitatifs. C'est d'ailleurs pour ces derniers que le procédé a été introduit dès le XVIII^{ème} siècle afin de représenter les trajectoires individuelles de voyageurs (Bahoken, 2010).

D'après B. Dent (1999), les cartes de flux les plus anciennes seraient celles réalisées par E. Halley en Angleterre en 1701, ou encore par P. Buache, en France en 1752, pour illustrer le mouvement de phénomènes atmosphériques à l'aide du procédé des isolignes (Dent, 1999 : 225), présenté sur la Figure 4-19.

¹⁴⁶ Certains travaux permettent d'ailleurs d'intégrer le son aux représentations, comme ceux de S. Caquart (2001).

¹⁴⁷ Seul le cas des figurés linéaires est présenté, les modifications apportées sur les autres types de figurés (ponctuels et surfaciques) n'étant pas spécifiques à la cartographie de flux : elles concernent le champ plus général de la cartographie statistique.

Figure 4-19. Exemple de carte réalisée par E. Halley (1701)



Edmund Halley's 1701 map of the known magnetic variations of the world.

Source: Dead Reckonings, Lost Arts in the Mathematical Science. <http://myreckonings.com/wordpress/> [lien vérifié le 06/08/2015]

Les flux issus de l'observation de processus économiques et sociaux sont plutôt représentés par des liens qui relient des couples de lieux d'origine et de destination. Plusieurs règles empiriques sont mises en œuvre pour les représenter :

- un trait droit est tracé entre deux lieux qui sont localisés dans un espace, généralement euclidien ;
- la *longueur* du trait correspond à l'éloignement des lieux : plus précisément à une acception de la métrique concernée, généralement la distance euclidienne (ce qui explique aussi le fait que le trait soit droit) ;
- la *largeur* éventuelle du trait correspond à la valeur du flux qu'il représente ;
- la *forme* du trait est rectilinéaire, parfois curvilinéaire pour des raisons esthétiques ;
- ce trait rectilinéaire peut être tracé en deux voire trois dimensions ; il peut être complété, par une pointe lorsque le flux est orienté et dans ce cas, il correspond alors à une flèche.

Différents auteurs ont suggéré des variantes à cette formalisation graphique générale du trait droit. Elles se distinguent selon qu'elles modifient le *Grain* (ou la texture du trait), ou ses dimensions géométriques, sa taille lors de la mise en œuvre de la variable visuelle *Taille*.

4.2.1.2.1 Le Grain du trait

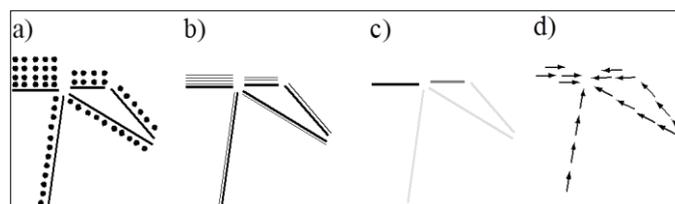
Dans le chapitre 12 de son ouvrage *Cartography: Thematic Map Design*¹⁴⁸ consacré à la conception des cartes de flux, B. Bent (1999, 2012) souligne la pléthore de formes autres que la flèche qui peuvent être mis en œuvre pour représenter des flux.

L'auteur s'appuie pour cela sur les résultats de l'analyse réalisée par M. J. Park (1987) d'une centaine de cartes des flux publiées dans 71 manuels scolaires, entre 1891 et 1984. Trois types de cartes sont distinguées (Park, 1987 : 50-66) : la carte radiale (ou carte en oursins) ; la carte en réseau (des flux planaires) et la carte de flux statistiques au sens où nous l'entendons. B. Dent suggère aussi de faire varier le style du lien, en utilisant des symboles graphiques qui ne sont pas couramment mis en œuvre

¹⁴⁸ Une version en PDF de ce chapitre est disponible à l'adresse suivante : <http://gis.depaul.edu/shwang/teaching/geog360/private/FlowMap.pdf> [lien vérifié le 02/08/2015].

pour illustrer des flux OD. Il présente, en particulier, quatre variantes au trait droit qui sont illustrés sur la Figure 4-20.

Figure 4-20. Illustration d'une série de variantes autour du trait droit, proposées par B. Dent (1999)



Source : D'après B. Dent, 1999.

Les semis de points régulier et les lignes parallèles (voir Figure 4-20 a) et b) consistent à appliquer respectivement, un nombre de points ou de lignes proportionnel à la valeur du flux, comme si l'on appliquait un *Grain*. Ces propositions ne s'appliquent toutefois qu'à la représentation de valeurs de flux qui sont uni latérales, issues de matrices symétriques, de la même manière que la teinte mobilisée sous la forme d'un aplat. Le résultat obtenu conduit à un figuré dont la largeur est proportionnelle à la valeur du flux, afin de conserver le rapport de proportionnalité des valeurs (Dent, 1999 : 229). A noter qu'il aurait pu être intéressant de conserver l'idée de densité de points ou de lignes pour illustrer la part du flux représenté par le lien, dans le total général, en recourant à la Valeur. Nous reviendrons sur cette possibilité dans la troisième section du chapitre 7. Les lignes colorées (voir Figure 4-20c) sont plus aisées à mettre en œuvre pour représenter des flux immatériels : elles correspondent à la combinaison des variables visuelles *Valeur* et *Couleur* qui permettent de représenter les flux en k classes de valeurs assorties d'une nuance de teinte.

Ces propositions intéressent surtout le champ de la cartographie de flux matériels (affectés sur réseaux), telle qu'elle est réalisée dans le champ des transports. Elles nous semblent peu applicables à la cartographie de matrices OD. C'est pourquoi nous les considérons comme secondaires, voire facultatives : la mise en œuvre de la *Taille* étant, à l'inverse, fondamentale dans la conception d'une carte de flux.

4.2.1.2.2 La largeur (et/ou la hauteur) du trait

La variable visuelle *Taille* a pour objet de traduire graphiquement la relation de proportionnalité entre la valeur du flux disponible dans la matrice et l'apparence du lien, autrement dit, ce que voit et perçoit l'observateur. Sa mise en œuvre consiste à faire varier l'une des dimensions du figuré : sa longueur, sa largeur, sa surface ou sa hauteur de manière proportionnelle à la valeur correspondante. Elle permet ainsi de rendre compte des différences de proportionnalité des figurés de manière ordonnée : les plus petits se distinguant des plus gros de manière hiérarchique. La *Taille* est ainsi également associative et sélective puisqu'elle rend compte de la proportionnalité de manière différenciée, sélective.

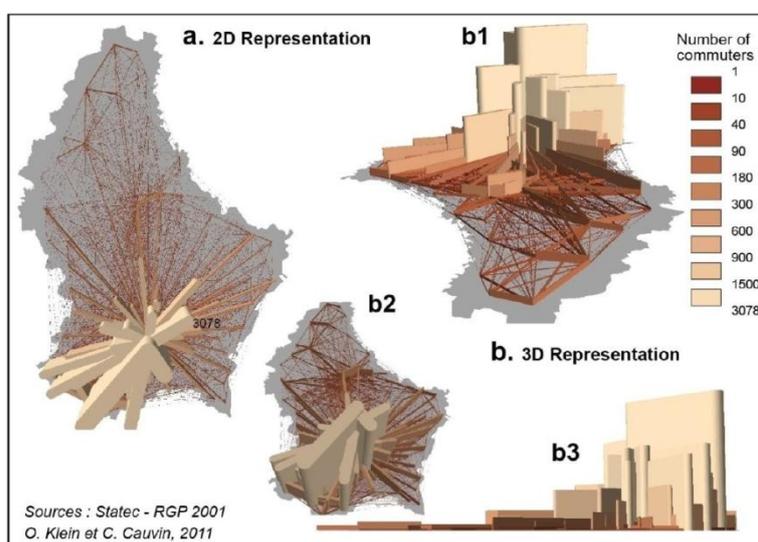
Etant donné que la position des lieux est fixe dans un système de projection donné, la mise en œuvre de la *Taille* en cartographie de flux consiste à faire varier la *largeur* du figuré, celle du trait ou de la bande et non sa surface. Les raisons en sont simples : d'une part, la *longueur* du trait est déjà mobilisée pour illustrer la distance (généralement euclidienne) parcourue par le flux, puisque ce dernier lie une paire de lieux ; nous y reviendrons dans la 3e section du chapitre 7. D'autre part, la mise en œuvre classique de la *Taille*, proportionnellement à la surface du figuré, en utilisant le rapport ($longueur * largeur$), est intenable en cartographie de flux, en raison de la composante spatiale du référentiel euclidien traditionnel (nous y reviendrons dans les chapitres 6 et 7).

Deux variantes à cette proposition classique d'application de la *Taille* ont été proposées. La première consiste à faire varier la *hauteur* de la bande et la seconde, à modifier la *longueur* du figuré, en fonction de la valeur du flux.

La solution qui consiste à faire varier la *hauteur* du figuré consiste à réaliser une carte en trois dimensions, où la composante (Z) de la donnée localisée est aussi assimilée à une altitude. Il s'agit de l'application aux données de flux d'une méthode de cartographie statistique suggérée dans les années 1960 qui consiste à transformer les données, puis à les considérer « [...] *simplement comme une surface statistique, où la hauteur (par exemple la densité de population rurale) varie à peu près de la même façon que l'altitude sur les cartes topographiques* » (Haggett, 1965 : 241). Cette suggestion a été explorée par différents auteurs (Tobler, 1987 ; Dorling, 1991 ; Klein et Cauvin, 2011).

La figure la plus récente de la carte de flux qui en résulte est illustrée par la Figure 4-21 réalisée par O. Klein et C. Cauvin.

Figure 4-21. Exemples de cartes de flux en trois dimensions, d'après O. Klein et C. Cauvin (2011)



Source : O. Klein et C. Cauvin, 2011.

Sur la Figure 4-21, les auteurs font varier deux dimensions du trait : sa *largeur* et sa *hauteur* proportionnellement à la même valeur de flux. Ce procédé correspond à une double application de la même variable attributaire (la valeur de flux) et visuelle (la *Taille*). La bande est par ailleurs segmentée en *k* classes d'effectifs présentant des nuances de teintes variables, par l'application conjointe des variables visuelles *Valeur* et *Couleur*. On notera également sur cette carte que les valeurs de flux les plus importantes sont assorties des nuances les plus claires et inversement. Ainsi, plus les valeurs de flux sont faibles, plus leur teinte est foncée, leur *largeur* et leur *hauteur* petite. Les auteurs agissent également sur la luminosité des bandes en trois dimensions, en introduisant une ombre qui rehausse leur discrimination visuelle, ces dernières n'ont donc pas besoin d'être assorties d'un contour. D'après les auteurs, cette carte met en évidence les « *contrastes* » de la carte de flux.

La solution qui consiste à faire varier la *longueur* du figuré correspond à une variante de la carte de flux réalisée en trois dimensions. Il s'agit des « *chronocartes* » ou « *Cartes en relief d'espace-temps* » proposées par A. L'Hostis (1996, 2003), que nous présentons dans la sous-section suivante.

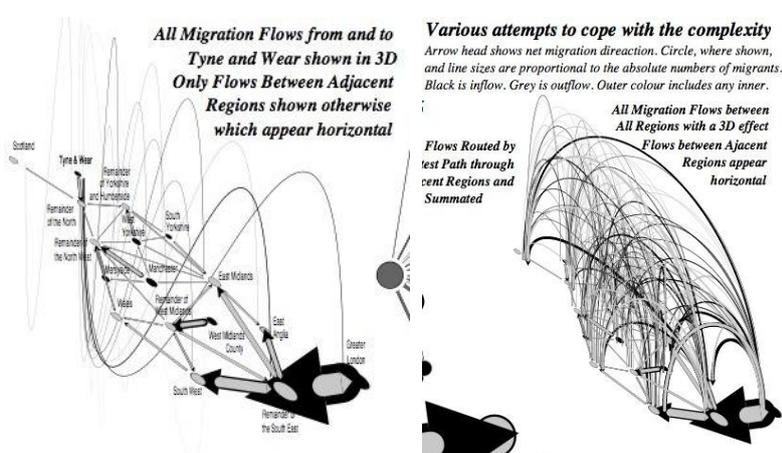
4.2.1.2.3 La longueur du trait

Il convient de noter au préalable que les cartes qui conduisent à la modification de la *longueur* du trait sont surtout expérimentales et assez confidentielles : elles n'ont pas connu une large diffusion qui aurait conduit à leur appropriation par les cartographes et les développeurs d'outils de cartographie. La raison tient probablement au fait qu'elles s'inscrivent dans le cadre de la théorie des graphes. Elles sont, en effet, issues de l'exploitation de graphes planaires, c'est-à-dire d'une matrice de liens (Lij) affectés sur un réseau typé (ou un multigraphe), où à chaque arête correspondent plusieurs modes de transport. La « *chronocarte* » de A. L'Hostis (1997, 2003) a ainsi l'apparence d'un graphe planaire où le trait est soit projeté sur la troisième dimension, soit déformé sur la *longueur*. Si elle ne s'apparente pas à une carte de flux au sens où nous l'appréhendons (le lien n'est pas valué), nous mentionnons toutefois l'idée mise en œuvre car elle est doublement intéressante : nous y reviendrons dans le chapitre 6 consacré aux solutions liées à la prise en compte de la distance parcourue par le flux.

La représentation de la distance parcourue par le flux s'inscrit, en effet, dans la résolution d'un double problème qui est à la fois théorique et d'ordre graphique. Sans rentrer dans le détail d'une analyse présentée dans le chapitre 6, nous présentons les deux familles de solutions qui ont été apportées, à savoir la modification de la longueur du trait de manière graphique ou mathématique.

Dans le premier cas, la méthode graphique consiste à appliquer un angle au trait pour le transformer en courbe. Il s'agit d'une approche très commune considérée d'ailleurs comme plus esthétique, elle préside d'ailleurs à la mise en œuvre des nouvelles manières de tracer les liens apportées par l'Informatique graphique que nous avons présentées dans la section 4.1.6, page 169. Dans certains cas, cette incurvation du tracé est combinée à une élévation, comme le fait D. Dorling (1991), dont une image est présentée sur la Figure 4-22.

Figure 4-22. Illustration d'une version du trait incurvé et en 3D

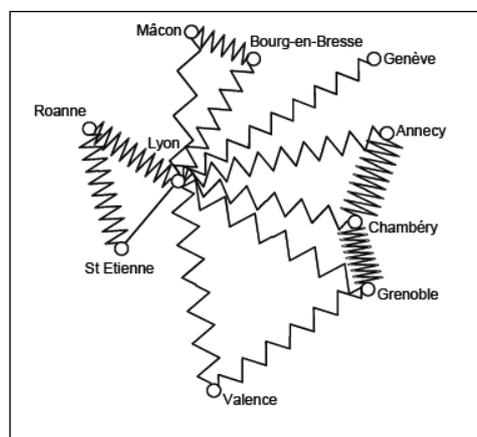


Source : D. Dorling, 1991

La seconde manière de modifier la *longueur* du lien est d'ordre mathématique. Elle agit sur la « *distance graphique* » perçue, en même temps que sur la distance euclidienne puisqu'elle affecte le caractère rectilinéaire du trait, toujours en tenant compte de la position des lieux. Cette approche suggérée par W. Tobler (1997) consiste à transformer le tracé du lien en substituant à la géométrie euclidienne de référence, une métrique réticulaire ou de transport, sans pour autant tomber dans l'anamorphose. L'auteur introduit pour cela la notion de ressort (*spring map*) qui consiste à transformer le trait droit du flux OD en un trait élastique, de manière à ajuster sa *longueur* par rapport à la distance parcourue par le flux sur le terrain.

La Figure 4-23 illustre ce principe de déformation du trait entre deux lieux pour aboutir à la réalisation de « cartes en ressort ».

Figure 4-23. Illustration du principe du « ressort », par A. L’Hostis (2003)



Source : A. L’Hostis, 2003.

D’après A. L’Hostis qui a examiné de près cette question liée à la représentation graphique de la distance sur réseau, le procédé du *ressort* suggéré par W. Tobler correspond à la notion de « longueur visuelle », différente de celle de « distance graphique » introduite par J.-C. Müller. L’auteur définit la « longueur visuelle » du lien, telle que perçue sur la carte, comme « [...] la longueur [réelle] du chemin correspondant. Si, par exemple, le chemin prend la forme de plusieurs segments mis bout à bout, la longueur visuelle correspond à la somme des longueurs des segments. La longueur visuelle n’est donc pas une notion euclidienne. Il faut remarquer que, définie de cette façon, la longueur visuelle n’est disponible entre deux points que s’il existe un chemin] (L’Hostis, 2003 : 219).

L’intérêt du *ressort* est donc amoindri par la difficulté de percevoir de manière instantanée la distance réelle correspondant à la *longueur* du lien. On voit bien sur la Figure 4-23 que c’est plutôt la multiplication de pics – entre des lieux qui sont proches ou pas – qui traduit l’importance de la distance parcourue, probablement le coût du déplacement compte tenu de certains critères (topographie, temps, coût financier, ...). De ce point de vue, la proportionnalité de la distance figurée par le tracé n’est donc perçue qu’indirectement et non instantanément. Il s’agit là d’une limite qui explique probablement la faible diffusion de ce procédé, au profit du plus classique trait droit, mis en œuvre dans l’espace euclidien traditionnel. Nous verrons dans le chapitre 6 qu’il est possible d’agir sur la perception directe et instantanée de la distance parcourue par le flux, en corrigeant plutôt la *largeur* du lien de manière à représenter la valeur non pas d’un flux mais d’un mouvement, en fonction d’une acception de la distance parcourue. Un autre biais de la « carte en ressort » concerne la représentation de la distance parcourue qui ne serait pas directement proportionnelle, mais laissée à l’appréciation de l’observateur de la carte. C’est notamment la raison pour laquelle A. L’Hostis proposa un autre formalisme de représentation des distances, dans le cadre des « cartes en relief d’espace-temps » ou « chronocartes » qu’il introduit dès 1996.

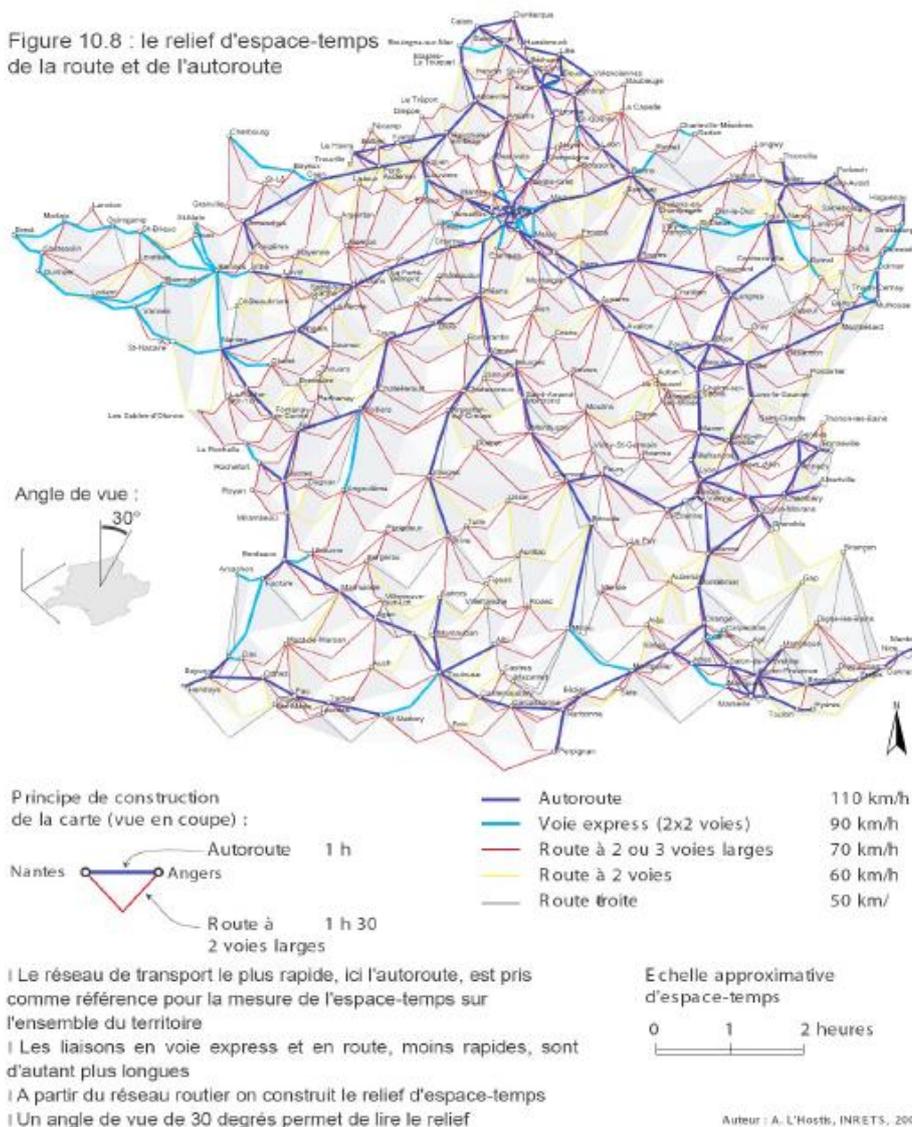
Les « chronocartes » correspondent à des figures dont l’objectif est d’intégrer « [...] la coexistence de moyens de transports aux performances diverses » (L’Hostis, 2003, 2007), c’est-à-dire une caractéristique sémantique du trait : le mode de transport, en l’occurrence les différents modes qui permettent de se déplacer entre un couple de lieux éventuellement la durée de déplacement. Leur particularité est de s’intéresser la représentation de la *longueur* du trait – qui caractérise donc la distance entre les lieux – de manière à prendre en compte l’hétérogénéité du déplacement sur un lien

donné, en fonction du mode de transport (autoroute *versus* route nationale). Etant donnée la forte hétérogénéité du réseau de transport liée à l'éventuelle multi modalité du déplacement, l'auteur mobilise la troisième dimension afin de représenter la distance du déplacement exprimée en temps de parcours (la distance-temps) de la manière suivante.

« La représentation des distances dans un réseau homogène avec des liaisons peu éloignées de la ligne droite est aisée : sur la réalisation dans le plan du graphe autoroutier [...], la longueur des arcs réalisés est directement proportionnelle aux distances-temps et les erreurs associées sont faibles. Par contre, dans un réseau hétérogène, dans lequel le différentiel des vitesses de parcours (ou de coûts) est important, la proportionnalité des longueurs des arcs ne peut être respectée si ceux-ci sont dessinés sous la forme de segments dans le plan. Dans les cartes en relief d'espace-temps, les arcs moins performants sont tracés sous le plan des nœuds sous la forme de deux arcs brisés, de telle façon que leur longueur soit proportionnelle aux durées de parcours » (L'Hostis, 2003 : 224).

Ainsi, la représentation de la distance apparaît déformée, par rapport à la géométrie euclidienne classique, par le recours à la troisième dimension et par la segmentation éventuelle du trait pour (éventuellement) absorber la durée du déplacement. La Figure 4-24 représente une « *chronocarte* » réalisée à l'échelle de la France, par A. L'Hostis.

Figure 4-24. Exemple de “chronocartes” (L'Hostis, 2000, 2003)



La variété des exemples de cartes de flux disponibles dans la littérature montre la richesse des cartes conçues avec ce trait droit ou légèrement incurvé, plus ou moins customisé, parfois orienté pour prendre la forme d'une flèche. Différents auteurs ont consacré des sections entières de leurs ouvrages (ou de leurs thèses) à la conception de la carte de flux (Dorling, 1991 ; Zanin et Trémélo, 2003 ; Dent, 1999). Tous s'accordent sur le fait qu'elle « [...] implique de résoudre de nombreuses difficultés » (Zanin, Trémélo, 2003 : 173) que nous appréhendons comme autant de contraintes liées à leur conception. Pour illustrer les différents problèmes liés à la mise en œuvre du trait droit, ou de la flèche, présentons d'abord la réalisation du tracé et les principes de sémiologie graphique y afférant puis les deux principales familles de variantes que nous avons pu identifier.

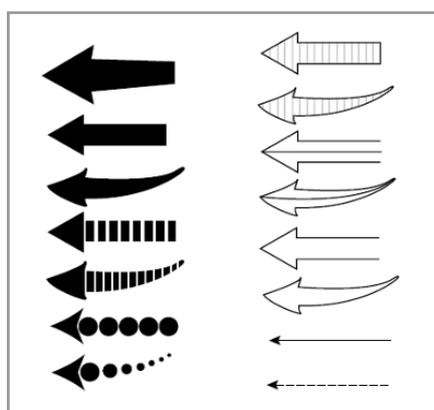
4.2.2 Du tracé à la disposition du figuré

La réalisation d'une carte de flux est généralement réalisée à l'aide de traits, de flèches ou de bandes dont la création et le placement sur la figure sont soumis à de nombreuses contraintes graphiques. Celles-ci portent sur leur conception même : sur le dessin de leur corps et/ou de leur pointe, sur la différenciation des lieux d'origine et de destination et sur la disposition d'ensemble. Il convient de noter que ces contraintes graphiques contribuent à renforcer *l'effet spaghetti*.

4.2.2.1 Le tracé du corps

Dans le cas de flux de données quantitatives, le corps de la flèche est nécessairement une surface mesurable sur laquelle pourra être appliquée une ou plusieurs variables visuelles : « *l'épaisseur, la forme ou la couleur du corps de la flèche, traduisent la nature, l'ordre ou la quantité de phénomène en mouvement* » (Zanin, Trémélo, 2003). Le tracé du corps prend donc la forme générale¹⁴⁹ d'une bande, orientée ou non (voir Figure 4-25).

Figure 4-25. Les types de flèches



Source : J. Bertin, 1973.

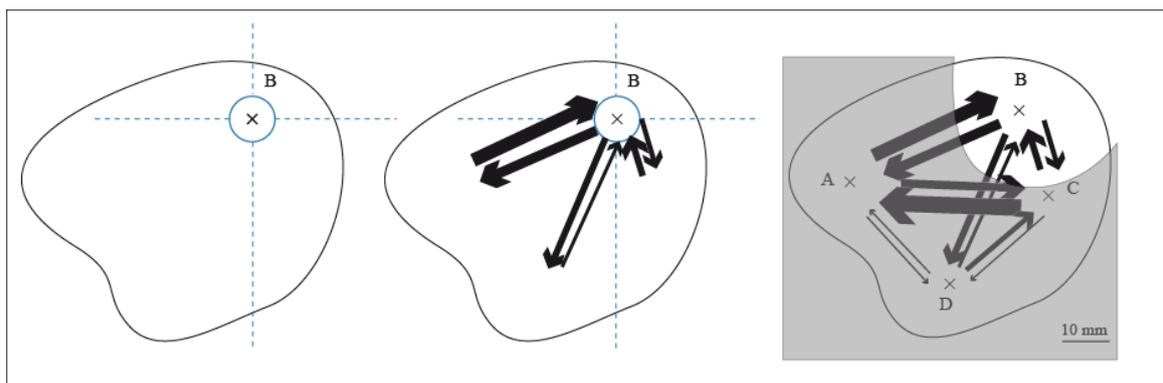
Le tracé du corps de la flèche varie, en pratique, en fonction du type de flux dont il est question. Il est, en effet, d'usage de représenter le corps de la flèche par une ligne pleine pour les flux de nature qualitative et par une ligne droite pleine pour les flux quantitatifs. Si le flux représenté est discontinu dans le temps, il convient alors de représenter le corps de la flèche par une ligne interrompue formée par un ensemble pointillés, afin de les distinguer des flux continus. Enfin, si la valeur d'un flux n'est pas connue avec précision et qu'il est nécessaire de le faire figurer sur la carte, le tracé correspondant

¹⁴⁹ D. Dorling (1991) a toutefois proposé un modèle de flèche hydride.

est alors illustré par une ligne droite, plus fine que les autres et surtout interrompu par un point d'interrogation : « ? ».

Le tracé du corps de la flèche varie également en fonction du type de flux, mais également de l'application de certaines conventions. La première d'entre elles consiste à centrer la flèche sur « l'axe du parcours ». Si celui-ci n'est pas précisé, comme dans le cas des flux immatériels OD, le principe consiste à placer le figuré sur une ligne virtuelle qui relie les centres des lieux d'origine et de destination afin de donner l'impression d'une convergence ou d'une divergence des flux vers ou depuis les lieux. Ce principe correspond à l'application du trait droit (voir Figure 4-26).

Figure 4-26. Illustration de l'effet de convergence (et de divergence) des flux



Lorsque les flux sont de nature quantitative, le tracé du corps de la flèche nécessite de représenter la valeur des flux ainsi que leur importance relative les uns par rapport aux autres. Pour ce faire, la largeur de la flèche doit être mesurable, potentiellement variable de manière proportionnelle à la valeur du flux, en application de la *Taille*. La tête de la flèche doit, quant à elle, être perpendiculaire à l'axe du corps sans l'étouffer : « [...] être simple et sa taille proportionnelle (en longueur et largeur) à l'épaisseur du corps » (Zanin, Trémélo, 2003).

La surface de la flèche peut également être utilisée pour représenter la trame du déplacement, en application du *Grain*. Celle-ci consiste à distinguer les flux matériels, qui nécessitent de s'appuyer sur une infrastructure physique (un réseau de transport, par exemple) et les flux immatériels. La notion de trame renvoie à la nécessité de différencier les « différences de nature des flux », ce qui passe par le choix d'une « variable visuelle exprimant la relation de différence ou d'association entre les données, en implantation linéaire : la forme, la couleur ou la texture ». (Zanin, Trémélo, 2003). Cette trame est portée sur une barre qui est, dans certains orientée, lorsque la direction du flux à un sens.

4.2.2.2 Le tracé de la pointe

L'orientation du flux ou sens du déplacement est une contrainte graphique très forte qui pèse sur le tracé de la pointe (ou tête) de la flèche. Elle consiste à illustrer le mouvement sur une figure qui est nécessairement statique et qui, de fait, nécessite l'utilisation d'un figuré statique donc immobile pour représenter un déplacement. Cela passe par l'apposition d'une pointe, au corps des bandes, qui indique le sens du flux. D. Dorling (1991) suggère que l'orientation de la pointe peut être marquée par une couleur différente de la bande, en fonction du sens du flux (aller ou retour).

Le tracé de la pointe a aussi pour objectif de représenter l'effet de convergence ou de divergence des flux, comme illustré par la Figure 4-26. C. Zanin et M.-L. Trémélo (2003) précisent à ce sujet que les pointes des flèches doivent arriver au centre des points d'arrivée et de départ, sans toutefois les masquer, en suivant les points cardinaux. Il s'agit, en pratique, de réaliser un détournement constant autour des points qui symbolisent les lieux, de réaliser une défoncée graphique qui permettra, notamment, de les distinguer en fonction de leur attractivité.

4.2.2.3 *La différenciation des lieux d'origine et de destination*

Le troisième groupe de contrainte graphique liée aux figurés porte sur la différenciation des lieux impliqués dans les échanges. La prise en compte des lieux entre lesquels se produisent les flux est, en effet, importante dans la réalisation d'une carte de flux. Ces lieux sont souvent représentés par des points, de *Taille* variable en fonction de leur attractivité ou de leur caractère répulsif. D'après D. Dorling, l'attractivité des lieux peut également être renforcée par le tracé d'une pointe proportionnelle à l'indicateur mesuré au niveau des lieux (Dorling, 1991 : 110)¹⁵⁰.

Cette variété de la taille des lieux a pour objectif de traduire l'ordre relatif des flux ainsi que la hiérarchie des valeurs, d'où le fait que soient souvent mises en œuvre de manière conjointe, la *Taille*, la *Valeur* et la *Couleur*. La prise en compte des différences d'attractivité des lieux, de l'effet de taille sur les flux consiste à représenter sur la même carte, les flux observés du point de vue des lieux d'origine et/ou de destination par des points, celle des flux réalisés entre les lieux avec des bandes.

4.2.2.4 *La disposition des figurés*

Le quatrième groupe de contraintes, et non le moindre, consiste à gérer la disposition des différents figurés sur la carte, selon qu'ils sont ponctuels ou linéaires. Lorsque ces derniers sont nombreux et présentent des tailles variables, un arbitrage est requis afin de sélectionner les figurés qui seront favorisés visuellement, en étant placés au premier plan. Cette sélection est généralement réalisée en fonction de la taille des figurés : de la largeur des liens ou de la surface des points. La pratique courante est celle du placement des plus petits points au premier plan, puis des flèches de manière à illustrer *l'effet de convergence* des flux vers des lieux et le *franchissement de la limite de zone*. Plusieurs contraintes sont, en effet, liées à la disposition des flèches sur une carte.

4.2.3 *Les contraintes de la flèche*

Les contraintes graphiques au placement de la flèche qui illustre les flux sont essentiellement liées à leur agencement sur le document, à l'espace disponible sur le support, à leur éventuelle orientation, leur taille mais aussi à leur échelle.

4.2.3.1 *Un problème de place et d'orientation*

Le problème de place des flèches résulte de l'une des composantes de l'effet de maillage lié à la position des lieux, qui impacte la grosseur des flèches. En raison de la première loi de la géographie de W. Tobler (1970), les flèches les plus grosses, celles qui illustrent les valeurs de flux les plus importantes, concernent généralement des zones qui sont proches ce qui pose un problème évident de place qui est lié au placement des figurés, et à leur chevauchement. Ce problème de place est surtout spécifique aux matrices de flux denses, caractéristiques de *l'effet spaghetti*. C'est la raison pour laquelle les solutions graphiques qui y sont traditionnellement apportées consistent à réduire le nombre de figurés, ou bien à éluder la flèche, en représentant les flux par une bande. Ces solutions présentées sur la Figure 4-27, illustrent notre matrice exemple.

Une première famille de solutions consiste à apposer un angle (α°) au tracé des flèches qui sont initialement rectilinéaires, afin de les transformer en courbes (de Beziers). Cette méthode purement graphique de déformation du figuré présente deux avantages : elle permet de représenter une quantité plus importante de figurés, en même temps qu'elle réduit leur chevauchement. La principale contrainte à sa mise en œuvre réside dans la conservation de l'équi répartition des figurés arrondis, en conservant

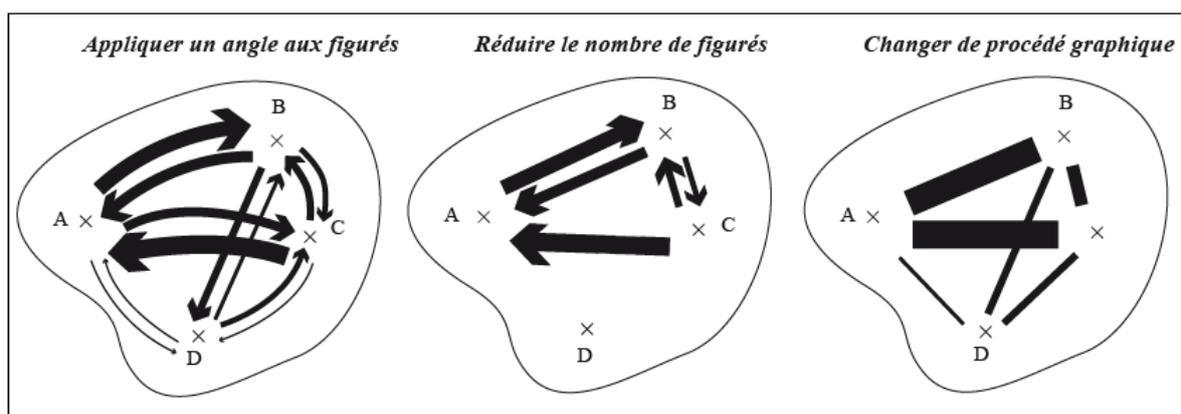
¹⁵⁰ Une illustration de ce procédé qui consiste à faire varier la pointe de la flèche de manière proportionnelle à la migration nette est disponible ici : http://www.sasi.group.shef.ac.uk/thesis/print_display.php?print=105 [lien vérifié le 02/08/2015].

une distance raisonnable de séparation visuelle, sans générer de chevauchements à d'autres endroits de la figure.

A noter que c'est une variante de ce procédé de déformation graphique qui prend la forme d'une fusion graphique, qui a été proposée par Phan et al. (2005) et D. Holten (2006) et, qui est par ailleurs à l'origine de la méthode de l'*edge bundling* que nous avons présentée dans la 4.1.6, située page 169. La seconde famille de solutions consiste à réduire la quantité de figurés à placer sur la carte, en appliquant un critère de sélection qui agit *in fine* sur la part d'objets. Nous y reviendrons dans les chapitres 8 et 9. La troisième famille de solution consiste à changer de procédé graphique, d'évincer la flèche au profit de la bande qui correspond à une flèche sans pointe, plus facile à manier. Elle est mise en œuvre lorsqu'il s'agit de représenter, certains types de flux mais pas d'autres. En effet, nous avons montré (chapitre 3) que la bande était adaptée à la cartographie des matrices de type (Fij^+) , décrivant un volume bilatéral.

Ces trois familles de solutions sémiologiques conduisant à réduire les problèmes de placement des figurés sont illustrées par la Figure 4-27, qui représente les flux de notre matrice exemple.

Figure 4-27. Illustration des problèmes de place des flèches et des solutions existantes



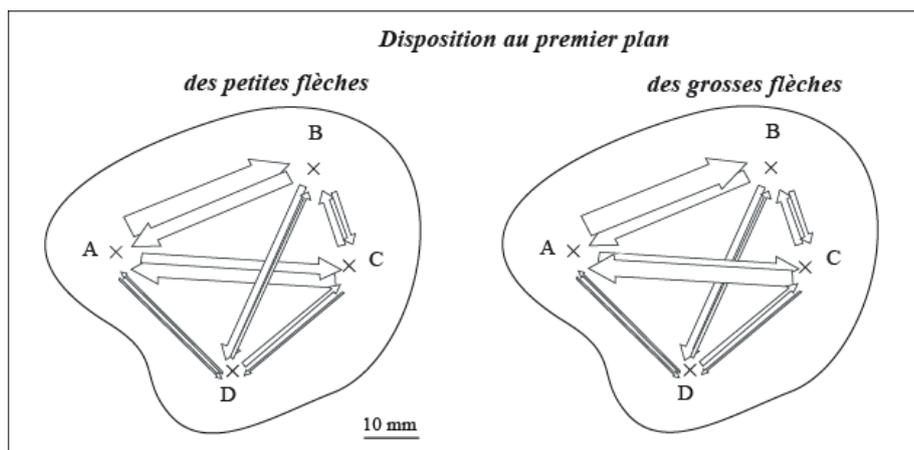
Il est important de noter que ces trois familles de solutions graphiques apportées au problème de place des figurés sont directement liées aux trois familles de méthodes de cartographie de flux, par conséquent au type de matrice. Elles dépendent de ce fait des éventuels traitements qui ont été réalisés sur les données, en amont de l'étape de représentation. Le changement de type figuré graphique (par exemple : le remplacement de la bande par une flèche) n'est rendu possible que parce que les valeurs de la matrice de données à représenter ont été modifiées, en l'occurrence que des valeurs de flux bilatérales (Fij) ont été décomposées en volume bilatéral (Fij^+) ou en solde bilatéral (Fij^-) .

4.2.3.2 Un problème de « taille » lié à la disposition des flèches

Il est d'usage très répandu de gérer la disposition des flèches en fonction de leur taille, en portant au premier plan celles dont les corps sont les plus fins, c'est-à-dire les petites flèches ; les grosses flèches sont alors, dans ce cas de figure, placées à l'arrière plan. L'explication tient à ce que W. Tobler qualifie de « *problème des lignes cachées* » qui réduisent la visibilité des petites flèches lorsqu'elles sont placées en arrière plan (Tobler, 1987 : 159).

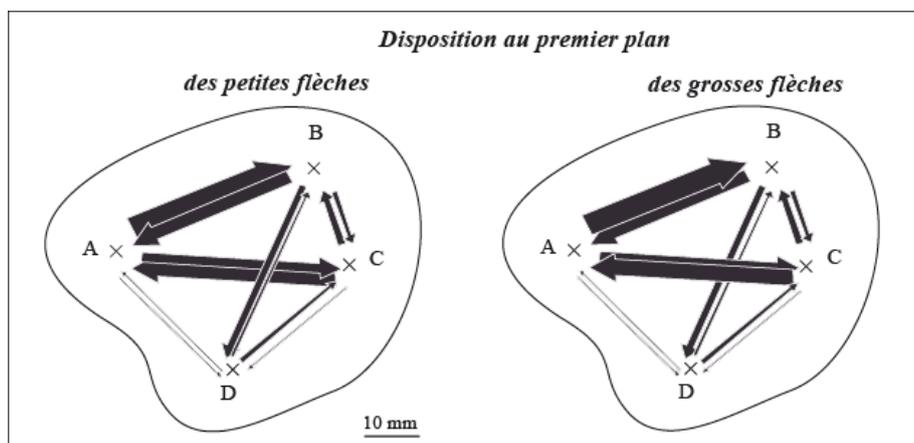
La Figure 4-28 illustre les deux possibilités de disposition des flèches, en fonction de leur taille, des valeurs de flux issues de notre matrice exemple.

Figure 4-28. Illustration des possibilités de disposition des flèches



La disposition des petites flèches au premier plan est aussi suggérée par C. Zanin et M-L. Trémélo (2003) de manière naturelle, alors qu'elle peut être contre-productive. L'un des objectifs de la cartographie de flux est, en effet, de donner à voir généralement les échanges les plus importants ce qui nécessite de placer les plus grosses flèches au premier plan. Placer les petites flèches au premier plan revient à privilégier les petits flux (en valeur) au détriment des plus gros flux qui sont illustrés par des grosses flèches. Au premier plan visible se trouvent alors représentés les flux dont la valeur est la plus faible. Si cette solution est efficace d'un point de vue visuel – puisque les plus gros flux restent quand même visibles sous les plus petits, en étant plus large – elle contribue à privilégier une contrainte graphique au détriment d'une contrainte méthodologique liée à l'affichage de la valeur du flux, alors qu'il devrait en être autrement.

C'est la raison pour laquelle W. Tobler a suggéré de placer ces petites flèches plutôt en arrière-plan, en justifiant sa suggestion par le fait que cette manière de procéder privilégie la visibilité des valeurs de flux les plus importantes (Tobler, 1987 : 159 ; Tobler, 2003 : 21-25). Elle offre aussi l'opportunité d'une lecture différente de la même information, par une réduction de l'encombrement de la carte, par conséquent de *l'effet-spaghetti*. Elle contribue toutefois à masquer une partie de l'information disponible puisque les petites flèches sont masquées. Pour résoudre ce « *problème des lignes cachées* » qui en découle, D. Dorling a suggéré, à la suite de la proposition de W. Tobler, d'appliquer le procédé de la *défonce graphique* au corps des flèches. Ce principe est illustré par la Figure 4-29, à partir des valeurs (Fij) de notre matrice exemple.

Figure 4-29. Illustration de la *défonce graphique* appliquée aux flèches

Couramment mise en œuvre en cartographie statistique pour représenter les points proportionnels, la méthode consiste à détourer les flèches avec un liseré blanc (Dorling, 1991 : 112). Les figurés, qui sont dès lors nécessairement teintés, voient ainsi leur séparation visuelle renforcée lorsqu'ils sont très nombreux. En cartographie statistique, la défonce est appliquée aux symboles ponctuels, lorsque les plus petits d'entre eux sont placés au premier plan, afin qu'ils puissent être distingués de ceux qui sont disposés en arrière plan. Elle devient inutile lorsque les symboles les plus petits sont placés en arrière plan, en raison de l'opacité de la teinte des figurés qui est généralement maximale. Ce principe, appliqué aux flèches, fonctionne de la même façon lorsque les plus petits flux sont placés au premier plan : dans ce cas, la défonce ne constitue pas un frein à la perception visuelle des plus petits et des grands flux, qui sont placés en arrière-plan. A l'inverse, lorsque les gros flux sont disposés au premier plan et que leur teinte est opaque, ils masquent par définition – en partie ou totalement – les plus petits flux, renforçant ainsi *l'effet de surcharge graphique*, caractéristique de *l'effet-spaghetti*. L'ajout d'une défonce présente alors un effet pervers car elle augmente la complexité de la figure en alourdissant les figurés.

Derrière la *surcharge graphique* trop importante est parfois associé le second effet graphique qui caractérise également *l'effet-spaghetti* à savoir *l'effet de couverture spatiale*. Nous avons montré (chapitre 2, section 2.3.3.1), que celui-ci est lié aux dimensions du figuré, en particulier à la *longueur* des flèches, laquelle est directement liée au maillage, à la forme des zones qu'il convient de franchir pour générer le flux.

4.2.3.3 Un problème de longueur, lié au franchissement de limite de zone

Le respect de la convention qui consiste à représenter un effet de convergence ou de divergence des flux (voir Figure 4-26) renforce le problème de place et celui du chevauchement des figurés. Il génère, en effet, un effet pervers lié au recours à la notion de barycentre des zones et à la longueur des figurés, celle-ci dépendant de l'éloignement des couples de lieux d'origine et de destination. C'est *l'effet de couverture spatiale*.

L'effet de couverture spatiale résulte de la nécessité de mettre en relation les centres des couples de lieux qui sont éloignés. Le trait correspondant traverse ainsi, dans bien des cas, tout ou partie de la zone d'étude masquant l'ensemble des informations situées sur son passage, qu'il s'agisse d'informations relatives au fond de carte ou aux tracés de longueur moindre. Deux solutions ont été apportées à ce problème.

La première solution, apportée par E.G. Ravenstein (1885), dans sa carte n°5 sur les courants migratoires (voir section 1), consiste à ne représenter que l'endroit du franchissement de la frontière sans toutefois relier les centres des lieux d'origine et de destination (Bahoken, 2012). Ce procédé donne l'apparence d'une continuité du phénomène tout en proposant une carte très lisible, l'essentiel des tracés ayant été supprimé, sans nuire à la transmission de l'information.

La seconde solution, suggérée par W. Tobler, vise à ne représenter que les flux qui ont lieu entre des zones contigües (adjacentes), qui partagent une même frontière (Tobler, 1987 : 162). Les autres flux, qui s'expriment à plus longue distance, sont ensuite réorientés vers les zones adjacentes. Ce procédé est une variante des flux affectés sur réseaux mais sous la forme de flèches. Il correspond à la représentation de flux sous la forme d'arbre hiérarchique qui a été rendue populaire ces dernières années par D. Phan et al. (2005), que nous avons présentée dans la section 4.1.6.1, page 170. Si la méthode contribue à aérer la figure, elle l'appauvrit considérablement dans le même temps puisqu'elle nécessite une sélection drastique de l'information disponible.

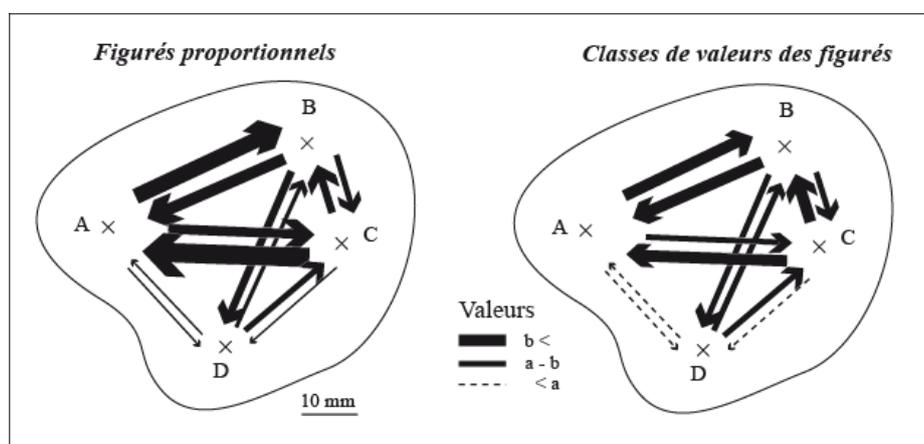
4.2.3.4 Un problème (de mise à) ou d'échelle des figurés

La question de la mise à l'échelle de la carte est sensible, car elle conditionne l'équilibre global de la figure. Cette question concerne d'une part, le choix du système de projection du fond de carte qui doit être adapté à l'échelle d'observation des flux représentés : une projection polaire apparaît ainsi plus adaptée à la cartographie de flux mondiaux qu'une projection de type Mercator. De l'autre, celle des figurés : la taille du symbole le plus petit doit rester visible sur la carte et celle du plus grand ne doit pas être trop envahissante.

La mise à l'échelle de la taille du figuré nécessite, par ailleurs, d'être cohérente avec l'amplitude de la distribution statistique des valeurs de flux, ce qui impose parfois de petits arrangements pour éviter de travestir la figure. B. Dent (1999) signale à ce sujet que, lorsque les plus grosses flèches sont trop larges, elles cessent d'apparaître comme porteuses des valeurs de flux les plus importantes, parce qu'elles apparaissent comme des symboles surfaciques et plus linéaires, ce qui est préjudiciable. L'auteur propose trois manières de résoudre ce problème de mise à l'échelle des figurés quantitatifs (Dent, 1999 : 228). La première solution consiste, étrangement, en cas d'impossibilité de mise à l'échelle adéquate des figurés, d'abandonner la carte et de recourir à un autre modèle de représentation des flux (voir chapitre 2). La seconde solution est plus classique puisqu'elle consiste à déterminer un seuil en dessous duquel les valeurs des flux ne seront pas représentées par des figurés de taille proportionnelle, et de symboliser les valeurs éliminées par une flèche de taille unique, représentée par exemple, par des pointillés. Cette solution revient à sélectionner les figurés qui seront présentés sur la figure, de la même manière que pour gérer les problèmes de place. La troisième solution consiste à créer k classes de valeurs de flux discrètes puis à déterminer une taille standard de figurés pour chacune des classes. Il s'agit d'une solution classique, très fréquemment mise en œuvre, qui consiste à ré-échantillonner les valeurs de flux de manière proportionnelle au milieu de la classe.

La Figure 4-30 illustre la procédure de mise à l'échelle des figurés, à partir des valeurs de notre matrice fictive (Fij). Cette procédure est une solution rapide et efficace, plus facile à mettre en œuvre que la règle générale de proportionnalité de la *largeur* des figurés avec la valeur qu'elles représentent. Il est important de noter la modification de la perception des différentes tailles des figurés, par conséquent des valeurs de flux sur les exemples de la Figure 4-30 : celles-ci apparaissent plus lissées, et les différences de valeur moins importantes. Nous reviendrons sur cette question de la perception des valeurs de flux traduite par les dimensions du figuré, dans le chapitre 6.

Figure 4-30. Illustration de la mise à l'échelle des figurés



Source : D'après B. Dent (1999), p. 228.

La flèche est, depuis toujours, privilégiée pour la représentation des flux. Sa mise en œuvre pratique pose cependant de nombreuses contraintes, qu'elle soit réalisée manuellement ou de manière automatisée puisque, dans bien des cas, le tracé automatisé est repris à la main. Rares sont les outils qui permettent de gérer simultanément les différents problèmes de conception graphique qui se posent. Le procédé de la flèche présente, par ailleurs, l'inconvénient de ne pas permettre la représentation simultanée de plus de deux variables. C'est pourquoi différents auteurs ont proposé l'utilisation de variantes au procédé de la flèche.

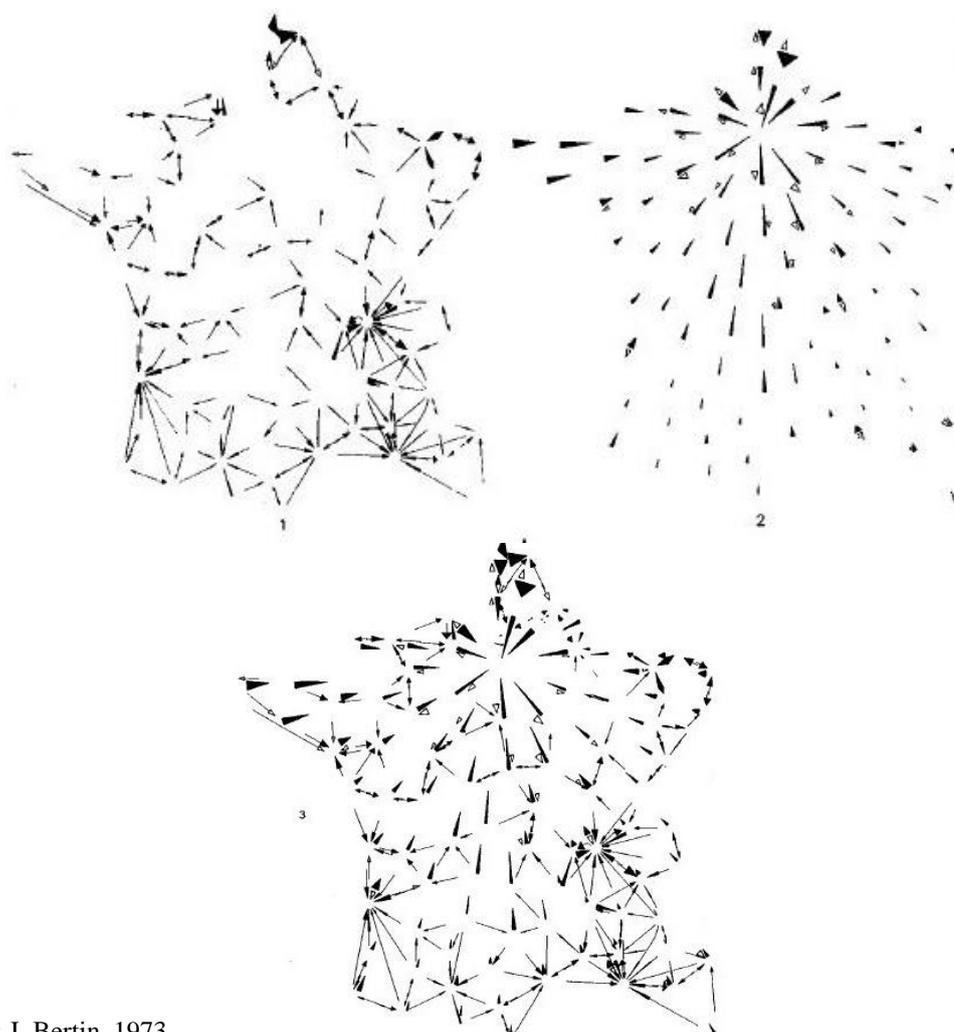
4.2.4 Les variantes de la flèche

Deux principaux groupes de variantes à la flèche peuvent être mentionnés : il s'agit de figurés fondés sur la pointe ou sur le corps de la flèche.

4.2.4.1 Le triangle

Le procédé de représentation des flux, à partir des seules pointes des flèches, est qualifié de triangle par J. Bertin, lequel semble l'avoir été mis en œuvre, pour la première fois, à notre connaissance, dans l'Atlas interdépartemental de la Datar paru au début des années 1970. Il s'agit d'une variante de la flèche intéressante, car la « *surface des flèches est proportionnelle aux quantités absolues* » (Bertin, 1973 : 351), comme l'illustre la Figure 4-31. Ce triangle présente aussi trois côtés qui sont susceptibles de porter une information chacun.

Figure 4-31. Cartographie de flux selon le procédé du triangle (Bertin, 1973)

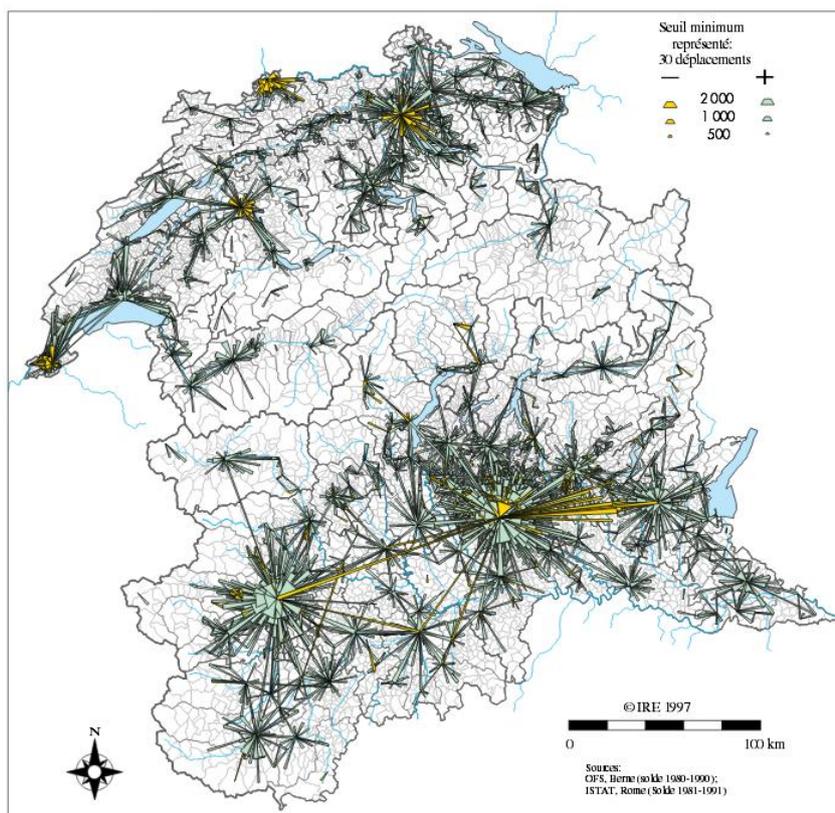


Source : J. Bertin, 1973.

Ce triangle, généralement de type isocèle, est parfois assorti d'un corps représenté par un trait droit. Le principe de construction du triangle est simple et efficace. Sa mise en œuvre est toutefois délicate, car il n'existe pas, à notre connaissance, d'outil disponible – sous un format libre ou propriétaire – permettant d'automatiser leur réalisation. Ainsi, sur une même carte de flux, tous les corps des triangles présentent des épaisseurs égales mais sont de *longueur* variable, car il est possible de faire varier leurs dimensions. J. Bertin décrit, comme suit, les cartes des migrations interdépartementales réalisées par S. Bonin : « *La longueur des flèches (triangle seul) est proportionnelle à la tendance migratoire (pourcentage d'émigrants) au dessus de 2%* » et leur disposition sur la carte : « *Les flèches se rapportant à la région parisienne (2) (Seine, Seine-et-Oise, Seine-et-Marne) sont situées dans les départements concernés (migrations vers Paris). Les flèches concernant les autres régions sont situées dans le département de destination (1)* » (Bertin, 1973 : 351). Dans la version originale de ces cartes, la pointe du triangle est placée au droit du lieu de destination pour signifier le sens du flux ; sa base est positionnée à une certaine distance de manière que ces deux éléments du triangle suivent une ligne virtuelle qui relie les centres des couples de lieux (Figure 4-31), la hauteur du triangle correspond à un indicateur du taux d'émigration et sa base (ou *largeur*) est proportionnelle à l'effectif de population née dans le département d'origine.

Il est important de noter que le triangle, tel que présenté sur la Figure 4-31 est le seul procédé qui autorise la mise en œuvre classique de la variable visuelle *Taille* pour représenter des flux, c'est-à-dire en utilisant un rapport de proportionnalité de la surface du figuré et non une seule de ses dimensions (généralement la *largeur*). Cette subtilité graphique présente un effet cognitif fort, en corrigeant le biais de perception de la répartition spatiale des valeurs de flux induit par le procédé de la flèche. Il réduit aussi sensiblement la perception de la distance parcourue par le flux puisque ces triangles, certes plus ou moins allongés (voir Figure 4-31 à droite), ne sont pas nécessairement assortis de corps identiques (voir Figure 4-31 à gauche). Enfin, l'utilité du procédé tient dans la description des flux s'exprimant entre des unités spatiales voisines, caractéristique d'un phénomène d'osmose (Bertin, 1973 : 351).

Une variante du triangle a été proposée à la fin des années 1990 par une équipe italienne, dans le cadre du projet *Cartogram*. Son principe de construction diffère nettement de la version proposée par J. Bertin (voir Figure 4-32). Le figuré est un triangle quelconque qui relie les couples de lieux : sa base est placée, comme dans la proposition de J. Bertin, au droit du lieu d'origine et sa pointe est placée au droit du lieu de destination. La base du triangle – et plus la surface – est rendue proportionnelle à la valeur du flux (Torriceli, 1997 : 9-13), ce qui signifie que la *Taille* est mise en œuvre de la même manière que la flèche, sur un côté du triangle. La surface du triangle exprime, quant à elle, la proportionnalité des effectifs de migrants concernés, soit un indicateur de la part du flux concerné par rapport au total de flux.

Figure 4-32. Cartographie de flux issue du projet *Cartogram* (Torricelli, 1997)

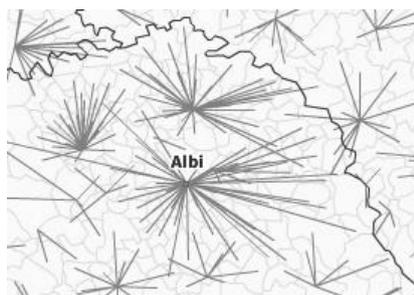
L'intérêt d'une représentation en triangle est double : lorsqu'il est mis en œuvre à la manière de J. Bertin, la surface du figuré varie en fonction de la valeur du flux. Le triangle permet également de renforcer l'effet visuel de convergence des flux. D'après Weger (1999 : 124), le procédé convient bien pour la représentation des « [...] phénomènes convergents ou rayonnants [...] par exemple : phénomènes de migration périodiques vers un petit nombre de destination (pour des raisons de lisibilité) [...] ». Le recours aux seules pointes de flèches se justifie aussi par le besoin de privilégier la destination des flux, tout en conservant l'allure générale de la carte en termes d'orientation. Il en résulte une carte présentant une série de triangles. Cependant, le procédé ne semble efficace que lorsque la quantité de triangles est faible, il ne permet donc pas de réduire à lui seul *l'effet spaghetti*.

Une seconde variante intéressante au modèle de la flèche consiste à supprimer ses pointes, en ne conservant que ses corps, c'est le procédé de l'oursin.

4.2.4.2 L'oursin

L'oursin est un procédé graphique très répandu dans un milieu où la cartographie de flux est surtout appréhendée en termes d'accessibilité à certains lieux centraux. Pour preuve, la facilité de sa mise en œuvre automatisée dans les outils existants, dédiés au géomarketing.

L'oursin présente deux caractéristiques principales qui sont illustrées sur la Figure 4-33.

Figure 4-33. Illustration du procédé de l'oursin

Source : Geoclip.fr¹⁵¹.

L'oursin n'est formé que des corps d'une sélection de flèches, car il n'a pas vocation à illustrer l'ensemble des liaisons entre tous les lieux. La valeur des flux et leur orientation sont par conséquent exclues *a priori* de la figure.

L'oursin correspond à un parti pris très fort de représentation qui doit être justifié par l'objectif de la carte. Sur la figure précédente, par exemple, la lecture est focalisée sur les flux polarisés par un lieu (A) depuis plusieurs zones de destination. Cette représentation se justifie bien pour la cartographie de matrices d'accessibilité, ou lorsque les zones d'origine ne sont pas nécessairement celles de destination. C'est le cas, par exemple, des zones d'emploi dont l'attractivité en termes de flux est observée depuis un ou plusieurs lieux d'origine qui correspondent aux lieux de résidence des actifs.

La seconde particularité de l'oursin est de ne représenter qu'une sélection de liens, sans tenir compte de leur valeur, dans l'ensemble. La carte de flux qui en résulte est alors très expressive : on devine bien l'orientation des liens par la structure d'ensemble des flux ainsi représentés ainsi que les lieux attractifs. Elle est cependant appauvrie, en l'absence d'information sur la valeur des flux. C'est pourquoi nous n'allons pas reprendre ce procédé de l'oursin, dans les chapitres suivants.

¹⁵¹ La version originale de cette image, en couleurs, est disponible à l'adresse suivante : http://www.geoclip.fr/aide/fr/fr/rubriques/autres_modes.htm [Lien vérifié le 02/08/2015].

Conclusion du chapitre

Les contraintes méthodologiques et techniques à prendre en compte pour cartographier une matrice de flux constituent autant de problèmes graphiques à résoudre au niveau des figurés. La conception d'une carte de flux nécessite aussi le respect d'une part, des principes de sémiologie graphique complétés de règles d'ensemble liées à la mise en page générale de la figure et d'autre part, de concevoir une légende qui soit efficace.

Les aspects liés à l'équilibre général de la mise en page de la carte sont généraux à l'ensemble des cartes statistiques, certains sont toutefois spécifiques à la représentation des flux. B. Dent (1999) insiste, par exemple, sur la nécessité de hiérarchiser les informations présentes sur la carte : une opération d'autant plus difficile que les figurés qui illustrent les flux sont particulièrement envahissants et qu'ils doivent être placés¹⁵² au premier plan de la figure et c'est bien là le problème. Certains figurés sont susceptibles de concerner plusieurs niveaux de lecture de la carte, de se chevaucher et d'entraîner des conflits visuels dont l'articulation conduit à l'effet spaghetti. L'usage de la transparence permet, par exemple, de gérer ces conflits sur le plan graphique, tout comme l'astuce du détournement proposée par (Dorling, 1991 : 114). Dans le cas de flux représentés à l'échelle mondiale, B. Dent (1999 : 227) évoque le problème engendré par la présence de figurés qui traversent l'ensemble de la carte, pour relier par exemple l'Asie du Sud à l'Europe de l'Ouest. Il s'agit là d'un problème lié au franchissement d'espaces continentaux et/ou maritimes qui concerne les aspects liés au graphisme des figurés ainsi que ceux, plus spécifiques, liés aux méthodes de cartographie à l'échelle mondiale, notamment ceux qui portent sur le choix du système de projection (Bahoken, Bauguitte, Lhomme, 2013).

Ces difficultés sont renforcées par le fait que « *Le caractère multidimensionnel des données [qui] entraîne des problèmes de conception cartographique qui sont inhabituels et complexes à résoudre* » (Dent, 1999 : 228) : la perception des flux doit être dominante et les figurés bien séparés visuellement. La toponymie et l'étiquetage éventuel des figurés ne doivent pas interférer avec l'intégrité visuelle du symbole. Les poncifs¹⁵³ et les motifs éventuels ne doivent pas engendrer de confusion avec les autres objets géographiques présents sur la carte. La mise à l'échelle des figurés doit autoriser une bonne lisibilité des plus petits et des plus grands (qui ne doivent pas être envahissants) symboles linéaires.

La représentation des matrices sous la forme de carte de flux suppose donc des choix dans les procédés de représentations (liens, flèches, triangles...) mais aussi dans les éléments qui composent la légende¹⁵⁴ : sur le plan graphique, S. Christophe (2009) en a proposé une synthèse récente ainsi que sur ceux qui résident dans les traitements de données réalisés en amont de la phase de cartographie.

Pour autant, si la légende doit être claire et permettre de faire le lien direct entre le dessin des figurés et les traitements des données réalisés en amont, la difficulté réside dans le choix de la méthode à appliquer, dans le type de raisonnement à mener pour aboutir à une cartographie efficace de la matrice, ce qui fait l'objet de la seconde partie de la thèse.

Les procédés de représentation cartographique des flux que nous avons présentés dans ce chapitre correspondent aux solutions classiques de sémiologie graphique. Les procédés sont mis en œuvre lors de l'étape de conception de la carte qui s'inscrit dans le cadre plus général de l'application de l'une des trois familles de méthode de cartographie de flux que nous avons présentées dans le Chapitre 3. Les figurés sont, par conséquent mobilisés, lors de l'étape de cartographie qui intervient, en pratique,

¹⁵² Au premier plan de la figure, par rapport aux autres éléments présents sur le fond de carte. Il s'agit d'une information qui doit « sauter aux yeux » d'après C. Zanin.

¹⁵³ Le poncif correspond à la répétition régulière d'un même figuré.

¹⁵⁴ Il s'agit là d'une suggestion de C. Zanin, lors d'une réunion de travail, en octobre 2012.

après celle d'identification du type de matrice, du type de donnée dont il est question mais également de l'objectif de la carte, du message à transmettre et du public auquel elle s'adresse.

A ce stade et compte tenu des apports méthodologiques et techniques liées au développement de l'informatique graphique dont nous avons fait état dans la section 4.1.6, il est possible de renouveler la figure des flux en agissant au niveau graphique.

Le chapitre 5 propose ainsi des solutions qui agissent sur la sémiologie cartographique des flux.

CONCLUSION DE LA PREMIERE PARTIE

« *C'est au dessin de la flèche qu'on reconnaît le bon dessinateur, celui qui a conscience d'apporter une réponse visuelle à tous les niveaux de lecture* ».

Jacques Bertin, 1973.

L'objectif de cette première partie était d'énoncer les problèmes, difficultés d'ordre méthodologique et contraintes graphiques liées à la représentation cartographique d'une matrice de flux. Quatre familles de contraintes ont pu être identifiées. La première, d'ordre théorique et conceptuelle, concerne la signification de l'information dont il est question ; la seconde est d'ordre méthodologique et analytique, elle porte sur la manière de représenter cette information que ce soit sur le plan graphique ou numérique ; la troisième famille est d'ordre visuel et porte sur la complexité de la figure ; la quatrième est de nature technique, elle est relative aux outils et programmes qu'il conviendrait de mettre en œuvre pour faciliter la réalisation de ces cartes de flux.

Cet objectif nous a conduits à définir, dans un premier temps, la matrice de flux devant faire l'objet d'une carte (chapitre 1). L'examen de la manière dont ses valeurs ont été obtenues a révélé une première famille de difficultés théoriques et conceptuelles, liées à la mesure, à la définition du type de flux. La mise en œuvre d'une approche compréhensive des notions et concepts mobilisés nous a conduits à souligner l'importance du dispositif de mesure utilisé, en particulier le rôle joué par les découpages temporels et territoriaux utilisés. Les problèmes qui en découlent ne sont pas inhérents à la cartographie de la matrice, ils la précèdent et doivent être résolus en amont (Bunge, 1966). Etant donné qu'ils n'influent pas sur la conception cartographique *stricto sensu*, nous n'allons pas nous y attarder. Nous allons supposer, par la suite de la thèse, que nous disposons d'une matrice complète, de données de flux observées mobilisant de nombreux lieux (plus d'une vingtaine). Cette matrice est notée (Fij).

La matrice ainsi définie, sa représentation graphique a été envisagée sous la forme de matrice ordonnée, de graphe ou de carte (chapitre 2). L'examen de ces trois modèles complémentaires et de leurs variantes nous a conduits à définir *l'effet spaghetti*, issu de la conjonction d'un effet de *surcharge graphique*, correspondant à la densité de teinte, et d'un *effet de couverture spatiale* lié aux dimensions géométriques des figurés. La nécessité de représenter l'orientation de certaines matrices traduit la contrainte de représenter *l'effet de convergence des flux*, déjà mentionnée par J. Bertin (1973). Cet effet est renforcé par *l'effet de taille des lieux*, une caractéristique de l'attractivité de lieux polarisant les flux les plus forts, sur de courtes distances, et définissant un système hiérarchique sur lequel nous reviendrons dans le chapitre 9. La contrainte qui se pose alors est celle du manque de place permettant de disposer les différents figurés, en raison de la nécessité de cartographier les flux selon une dialectique liens-nœuds, c'est-à-dire en mobilisant des figurés linéaires et ponctuels.

Le choix des figurés n'est cependant pas anodin, il dépend du type de matrice. Il convient, tout d'abord, de résoudre *l'effet d'amalgame* qui en dépend. Ce dernier est défini comme le mélange d'informations de nature différente, de ce qui relève ou non de la dynamique du système de flux (il est flagrant pour les matrices asymétriques). Il est résolu par la décomposition de la matrice en composantes (symétriques ou antisymétriques), ayant pour effet d'agir sur le choix du procédé. Son pendant graphique, *l'aspect compact* de la figure, est résolu par cette décomposition, mettant alors en évidence le *phénomène d'osmose* évoqué par J. Bertin (1973), autrement dit, l'importance des mobilités qui s'expriment entre des zones adjacentes, voisines ou des flux qui s'expriment à courte distance.

Dans les exemples de cartes que J. Bertin propose en 1973 (voir Figure 4.16, page 172), l'auteur représente des flux (Fij) et leurs relations (Lij) non orientées à une échelle régionale. En agissant ainsi, cet auteur illustre deux des quatre familles de méthodes de cartographie des matrices de flux (chapitre 3), l'ensemble étant validé par la rétrospective de l'image des flux présentée (chapitre 4). Cette rétrospective souligne aussi deux éléments notables dans l'évolution des solutions cartographiques apportées. D'une part, l'existence d'un lien ténu entre l'émergence des procédés cartographiques mis en oeuvre et les avancées théoriques et conceptuelles ; d'autre part, l'apparition progressive de solutions aux difficultés rencontrées, souvent en lien avec le développement technique et technologique.

Nous avons ainsi pu observer la manière dont différents problèmes ont été gérés, voire résolus, en fonction de la résolution de la carte, et au cours du temps.

A l'échelle mondiale, par exemple, *l'effet d'alignement topologique (ou effet de position) des lieux*, empêchant la perception claire des motifs des flux à cette échelle, est notamment résolu par le choix d'un système de projection approprié, dès les années 1920. Il l'était déjà par l'adoption du procédé distributif de Minard, *l'effet d'itinéraire*, lié à l'alternance terres-mers entre les couples de lieux. A l'échelle locale, les propositions portent plutôt sur le changement de modèle graphique. Elles conduisent, par exemple, à modifier la position des lieux.

L'évolution des solutions apportées au cours du temps est symptomatique des intérêts scientifiques et des moyens techniques de l'époque. La résolution de la contrainte liée à la *disposition des figurés*, par exemple, dont on sait qu'elle intéresse aussi la réduction de *l'effet de couverture spatiale* (en particulier la gestion de la superposition des figurés) est un exemple de cette évolution.

Dans un premier temps, ce problème de *disposition* est géré par un nouveau procédé : la forme distributive des cartes figuratives de C.J. Minard (1845). Avec l'avènement de la flèche, l'intérêt est davantage porté sur la sélection de l'information de flux à représenter, selon une focale particulière (les seuls lieux de destination) ou en fonction des valeurs de flux à représenter, les plus importants, en raison du formalisme du trait droit qui génère un chevauchement important des liens. Le développement de l'informatique, en particulier des Systèmes d'information géographique (SIG), entrainera l'apparition d'une vision continue des flux et des mouvements et des nouvelles images qui l'accompagne (telles que les cartes de flux financiers proposées par W. Tobler, 1981 a). L'approche discrète ne sera pas en reste, W. Tobler proposera les premiers outils permettant de tracer des flux sur une carte, de les sélectionner et de les préparer quand, dans le même temps, J. Bertin rassemblera les principes de sémiologie cartographique qui leur sont dédiés. Au fur et à mesure que la technique progresse, que l'informatique graphique se développe, le problème de la *disposition des figurés* se voit résolu par l'introduction de solutions graphiques inédites : le triangle de J. Bertin (1973) et, bien plus tard, la fusion graphique des liens de D. Holten (2006). L'introduction plus récente d'une nouvelle variable visuelle, la *Transparence* par MacEachren (1995), précédemment suggérée par J. Bertin (1973) conduira à annihiler définitivement cette contrainte. Cependant, c'est sans compter sur

l'émergence d'un effet graphique caché révélé par l'outil : la *profondeur des flux*, une variante de l'*effet d'amalgame* (nous y reviendrons dans le chapitre 5).

On observe ainsi qu'un problème donné, lorsqu'il est récurrent, peut être résolu en adoptant des méthodes variées, liées aux aspects cartographiques, au traitement des données numériques (mathématiques,...) ou à leur sémiologie ; les solutions d'un même problème pouvant être spécifiques à une contrainte particulière, ou communes à plusieurs d'entre elles. Compte tenu de la variété des problèmes et des solutions impliquées, nous avons fait le choix de les regrouper en quatre catégories définies comme suit.

Les difficultés d'ordre théorique et conceptuel intéressent la définition du type de flux dont il est question, ainsi que la construction des valeurs de la matrice (chapitres 1 et 3). Les difficultés cartographiques concernent, d'une part, la préparation du fond de carte (incluant le choix d'un système de projection et d'implantation des objets spatiaux) et la sélection des procédés de représentation ; d'autre part, les contraintes graphiques liées à la conception d'ensemble de la carte, au tracé des symboles et des figurés (chapitre 3 et 4). Les contraintes d'ordre sémiologiques portent sur le choix des variables visuelles et l'esthétique d'ensemble de la figure devant conduire à l'efficacité du message transmis (chapitre 4). Enfin, les contraintes d'ordre technique et technologique sont liées au support (papier ou numérique) et aux outils utilisés pour cartographier les valeurs de flux : les tracer sur un plan et les analyser (les résumer, les transformer...) et les visualiser.

Ces quatre familles de contraintes se manifestent à deux moments du processus de cartographie : lors de la préparation des données numériques et cartographiques ; au moment de leur représentation (aspect méthodologique) ou lors de la visualisation du résultat (aspect visuel). Le Tableau C-1 récapitule les principales contraintes et difficultés méthodologiques que nous avons pu identifier au cours des quatre premiers chapitres.

Tableau C-1. Récapitulatif des problèmes et des solutions envisageables

INTITULE (problème du à un effet)	CONTRAINTE	COMPLEXITE		SOLUTION			
		graphique	analytique	mathématique	cartographique		sémiologique cartographique
					liens	lieux	
spaghetti	visuelle	X	X	X	X	X	X
couverture spatiale	visuelle numérique	X	X	X	X		X
surcharge graphique	Visuelle numérique	X			X	X	X
Echelle (résolution)	méthodologique				X	X	
amalgame	numérique		X	X			
profondeur	numérique graphique		X	X	X		X
taille (des lieux)	numérique graphique	X	X			X	X
disposition	méthode visuelle				X	X	
convergence	visuelle	X			X	X	
position	visuelle numérique					X	
alignement	visuelle		X			X	
itinéraire	visuelle	X			X		

Dans le Tableau C-1, les différentes colonnes correspondent, de gauche à droite : au problème tel que mentionné dans le texte ; à la contrainte qui en découle, une difficulté méthodologique (lié à la préparation des données, au choix des figurés, ...) ou une contrainte graphique (le constat d'un nombre trop important de figurés) ; au type de complexité : graphique (au niveau du tracé des symboles, des figurés, de la position des lieux) ou analytique (concernant la définition de l'information de flux donnée et son interprétation) ; aux familles de solutions envisageables qui s'inscrivent dans des raisonnements mathématiques, cartographiques, sémiologiques ou plus largement liés au graphisme. L'attention est toutefois attirée sur le fait que les solutions techniques et technologiques apportées à ces problèmes ne sont pas présentées dans le Tableau C-1. Elles ne sauraient l'être car, comme nous l'avons déjà mentionné, la méthode précède l'outil qui intervient alors, nécessairement, dans un second temps de la réflexion.

Ces problèmes ont, dans l'ensemble, fait l'objet de solutions présentées dans cette première partie. Aussi allons-nous essayer de contribuer à résoudre, dans la seconde partie, une sélection d'entre eux permettant de réduire la complexité graphique de la carte de flux, surlignés en gris dans le Tableau C-1.

DEUXIEME PARTIE

Propositions de solutions visant à réduire la complexité de la carte de flux

INTRODUCTION DE LA DEUXIEME PARTIE

« L'expérimentation est nécessaire pour surmonter un grand nombre de problèmes de la cartographie de flux : si la question du 'Quoi ?' a été résolue, celle du 'Comment ?' reste à formaliser, dans le cadre d'un processus d'expérimentation qui devrait conduire à la détermination des meilleures – sinon des plus efficaces – méthodes cartographiques à appliquer aux données disponibles dans une matrice de flux ».

Daniel Dorling (1991), citant (Dale, 1971).

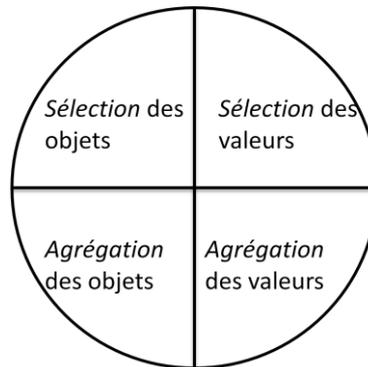
Cette seconde partie a pour objectif de présenter des solutions permettant de réduire la complexité graphique de la carte de flux. Celle-ci étant le résultat de la conjonction de différents problèmes identifiés dans la première partie, nous faisons le choix de contribuer à résoudre, sinon supprimer, les effets de *surcharge graphique*, de *couverture spatiale*, ainsi que *l'effet de disposition*, car ils impactent directement la perception des motifs de flux représentés sur la carte. Les autres effets que nous avons pu identifier font soit déjà l'objet de solutions que nous considérons comme opérantes, c'est le cas de *l'effet d'amalgame* résolu par une décomposition de la matrice (chapitre 3) ; soit mettent en œuvre des familles de solutions particulières qu'il n'est pas envisageable de présenter ici, faute de place.

L'effet de taille des lieux et de *convergence* des flux (vers les zones de plus forte attractivité), tout comme *l'effet d'alignement topologique* par exemple, peuvent être résolus dans le cadre de solutions liées aux transformations cartographiques de position (ou lors d'un traitement des données). Dans les deux cas, la solution apportée agit lors de la préparation des données numériques et/ou cartographiques, elle n'agit pas sur la construction de la carte de flux ni sur ses aspects sémiologiques. De la même façon, si la résolution de *l'effet de convergence* nécessite en pratique d'adapter le fond de carte à la taille des lieux (et/ou d'agir sur la topologie des lieux et/ou sur les dimensions des liens), de manière à ce qu'ils ne se superposent pas au droit des lieux d'origine et de destination, l'éventail des solutions envisageables est trop vaste, pour pouvoir être appréhendé ici.

Les solutions que nous proposons dans cette seconde partie ont une visée opérationnelle, celle de réaliser une carte lisible et compréhensible de la structure de flux représentés.

Nous avons vu dans la première partie que sa résolution constitue un véritable casse-tête qui a pu être résolu de différentes manières au cours du temps (chapitre 4), à l'aide de méthodes qui agissent généralement au niveau du traitement des données. Les solutions qui ont pu être apportées peuvent être regroupées en deux grandes familles qui intéressent, pour l'une les méthodes liées à la sélection et pour l'autre, à l'agrégation des objets géographiques ou des valeurs, comme illustré par la Figure I-2-1. Une troisième famille de méthodes peut être évoquée, il s'agit des méthodes complexes qui tiennent à la modélisation des données de flux et /ou aux transformations cartographiques des positions. Elles ne seront pas examinées, car elles ne contribuent pas nécessairement à réduire la complexité graphique de la figure (ce n'est pas leur vocation). Certaines d'entre elles visent d'ailleurs à introduire de la complexité visuelle, comme dans le cas des anamorphoses.

Figure I-2-1. Familles de solutions actuelles au problème de la complexité graphique de la cartographie de flux



Les solutions liées à la sélection des objets visent à réduire la quantité d'information à représenter sur la carte, celles qui portent sur la sélection des valeurs agissent comme un filtre des plus importantes ou des plus *significatives*. En effet, W. Bunge (1962, 1969) a proposé de résoudre la complexité graphique de la figure en la simplifiant, c'est-à-dire en sélectionnant avec « *parcimonie* » l'information à représenter. J. Bertin (1973) indique que « *Lorsque l'on veut représenter sur une même carte [tous les flux], des simplifications s'imposent. Elles peuvent être de différentes natures : 1° Représentation de la balance entre les deux sens [la composante (Fij)], 2° Ne pas représenter les mouvements au-dessous d'un certain pourcentage* » (Bertin, 1973 : 351).

Dans ce registre, nous verrons (chapitre 8) que l'approche la plus courante consiste à éliminer brutalement une très grande partie de l'information (sélection par les objets) afin de ne représenter que les flux qui sont considérés comme *significatifs* (sélection par les valeurs). Cela nous conduira à montrer que les deux approches communiquent et correspondent en fait à l'application d'un double critère de sélection. Ces méthodes de sélection sont-elles généralisables ? Sont-elles robustes ?

Les solutions liées à l'agrégation visent, quant à elle, à réduire la résolution des données. Lorsqu'elle est menée sur les objets, elle conduit, par exemple, à changer d'échelle (passer de la commune au département, des flux intercommunaux aux flux interdépartementaux) ce qui entraîne une réduction du nombre d'objets (de lieux), par conséquent du nombre de liens (d'objets) et d'une part de l'information (Fij). Si ces méthodes sont intéressantes, elles sont cependant à manier avec précaution. Le changement d'échelle géographique conduisant, d'une part, à un changement d'objet d'étude et d'autre part, à un arbitrage quand à la méthode d'agrégation à adopter. Le champ d'application de cette famille de solutions étant important, nous limiterons notre propos (chapitre 7) à l'approche structurelle.

Une troisième famille de solutions peut être évoquée, elle est liée à l'application de méthodes complexes qui conduisent à agir sur la localisation des couples de lieux (i,j). Ces méthodes sont également importantes, dans la mesure où elles permettent de résoudre des problèmes spécifiques liés aux *effets de position* des lieux (*effet d'alignement topologique*) Cette famille de solutions conduit ainsi à transformer les positions des lieux, par un changement de projection cartographique à l'échelle mondiale : par un changement de mode de représentation à l'échelle locale, en utilisant, par exemple, un modèle circulaire conservant la cardinalité des lieux (Bonin, Bahoken Mermet, 2013).

Quelles sont, dans ce contexte, les solutions que nous pouvons apporter aux problèmes de la cartographie identifiés ? En quoi diffèrent-elles des approches mentionnées précédemment, déjà mobilisées pour réduire la complexité de la carte ?

Les solutions proposées dans cette seconde partie intéressent, pour sa dernière partie (chapitres 8 et 9), des verrous classiques de la cartographie de flux liés à l'application d'un critère de sélection. Elles

s'inscrivent alors dans un cadre d'analyse classique, d'ailleurs non spécifique à la cartographie de flux. La partie s'ouvre avec des solutions que nous pensons plus originales. Elles sont d'abord d'ordre sémiologique (chapitre 5), puis liées à la prise en compte de variables d'espacement des lieux, dans le processus de cartographie de flux (chapitre 7).

W. Bunge indique que les contraintes conceptuelles doivent être résolues indépendamment de l'étape de représentation graphique des données (Bunge, 1966). Cela correspond, dans notre cadre d'analyse, à l'étape de préparation des données (chapitre 3). Cet auteur a insisté sur l'importance des aspects mathématiques à prendre en compte pour révéler les processus spatiaux, aux côtés d'approches plus graphiques liées à la représentation des données. Il suggère, pour cela, l'élaboration de cadres méthodologiques de « *fertilisation croisée* » (Bunge, 1966 : 27) qui articulent les aspects graphiques aux aspects mathématiques, résolvant d'abord les problèmes d'ordre mathématique avant les autres. Les aspects statistiques intervenant dans un second temps de l'analyse : « [...] *trouver la structure spatiale des objets en interaction, qui ont souvent des dimensions différentes, et les placer autant que possible relativement les uns par rapport aux autres à la surface de la terre* » (Bunge, 1966 : 210). Nous allons procéder à l'inverse, examinant d'abord la sémiologie cartographique des flux (chapitre 5) avant le traitement des données (chapitre 7, 8 et 9). Cette inversion n'est cependant qu'apparente, car nous allons supposer que les données numériques et le fond de carte sont d'ores et déjà prêts pour l'analyse.

Dans le chapitre 5, les solutions proposées sont liées aux aspects visuels, sémiologiques et esthétiques de la carte de flux. Elles n'agissent donc pas sur les données. L. Jégou (2013) a récemment rappelé l'inscription de la cartographie thématique dans le champ de l'iconographie, une branche de l'histoire de l'art, mais aussi de l'art moderne au sens de P. Kandinsky, de Paul Klee et d'autres dont les œuvres s'inscrivent dans une réflexion sur « *l'esthétique formelle* » et sur les « *règles de composition* » qui régissent la représentation dans son ensemble. Cependant, les travaux réalisés dans ce domaine (Chesneau, 2006 ; Christophe, 2009 ; Christophe et Ruas, 2009 ; Christophe, 2012 ; Christophe et Zanin, 2011 ; Dhée, 2013) s'inscrivent en cartographie (statistique ou topographique) générale, ils ne concernent pas les flux¹⁵⁵.

Les solutions proposées ici permettent de renouveler la cartographie des grandes familles de matrices de flux identifiées (chapitre 3), en mobilisant principalement la couleur. Nous introduisons dans un premier temps la transparence suggérée par A. M. Eachren (1995), comme variable visuelle (d'où la notation : *Transparence*) correspondant au canal alpha de la teinte, pour gérer *l'effet de couverture spatiale* (la superposition des figurés). Nous transférons ainsi des principes empiriques issus de l'Informatique graphique et les appliquons à la cartographie statistique des différentes catégories de matrices de flux, l'exemple d'application est une matrice de type (Fij⁺). Dans un second temps, nous opérons un transfert du principe de l'opposition colorée au niveau des liens (Lij), et l'appliquons à une matrice de type (Fij⁻). L'observation des résultats obtenus nous conduit vers la mise en évidence d'un *effet de profondeur*, à l'émergence de l'hypothèse de **la perception de plusieurs classes de longueur (Lij) de figurés**. Ce constat soulève un paradoxe, lié à l'hypothèse **d'un décalage entre l'information de flux perçue et celle qui est représentée sur le fond de carte**. Cet étonnement nous conduit à revenir, dans le chapitre 6, aux fondements de la cartographie de flux.

Dans le chapitre 6, nous opérons un retour sur les aspects théoriques et conceptuels de la construction du figuré qui décrit le flux sur une carte. Ainsi, avant d'en arriver à éliminer une partie de l'information de flux, la question qui se pose est celle de savoir s'il est si évident que cela de

¹⁵⁵ L. Jégou va d'ailleurs jusqu'à indiquer que les efforts menés pour améliorer l'esthétique des travaux complexes (en présentant l'exemple d'une carte de flux réalisée par une journaliste graphique bien connue) évoquent davantage la scientificité / technicité du travail réalisé (Jégou, 2013 : 395) que leur esthétique.

construire les cartes de flux telles qu'elles le sont habituellement, c'est-à-dire en agissant sur une seule dimension du figuré, sa *largeur*, alors que l'œil humain en perçoit plusieurs ? La pratique classique de la cartographie de flux, telle que nous l'avons formalisée dans le chapitre 3, est-elle réellement pertinente ?

La résolution de cette interrogation, dont on comprend aisément le caractère fondamental pour notre travail, ne saurait être réalisée aux seuls niveaux sémiologique de la teinte et statistique de la sélection de l'information. Elle intéresse les questions liées à l'aspect cognitif de la carte de flux, c'est-à-dire à la manière dont les figurés (décrivant les flux) sont perçus par un observateur isolé et *a fortiori*, celle dont ils sont cartographiés. Cela nous conduit à formuler l'hypothèse de **la légitimité de la pratique usuelle, que nous démontrerons par un examen approfondi de la construction de la carte de flux**. Pour la valider, nous inscrivons notre démarche dans la grille d'analyse proposée par C. Grasland (2009) et assumons la position de W. Tobler (1969), pour représenter, à titre d'exemple, une matrice symétrique de type (F_{ij}^+) . Nous proposons ensuite, à notre tour un cadre sémiotique d'analyse de la construction cartographique du flux. L'examen du *signifiant* et du *signifié* de la carte de flux nous conduit à analyser une série de quatre hypothèses qui interrogent la perception du figuré de ce double point de vue. Nous révélons ainsi des éléments nouveaux quant à l'interprétation de la carte de flux, lesquels nous conduiront inéluctablement à mener une réflexion en termes de *flux* (F_{ij}) et de *mouvement* (M_{ij}), mais aussi de *distance cartographique* et de *distance fonctionnelle*, dans le cadre d'une réflexion sur la *longueur* du figuré qui correspond, en réalité, à la *distance réelle apparemment parcourue par le flux*.

Cette analyse nous conduira au résultat selon lequel si la construction de la carte de flux est invariante, quelle que soit la matrice, (F_{ij}) ou (M_{ij}), et la thématique dont il est question, la différence entre deux cartes de flux réside dans l'interprétation de l'information représentée, c'est-à-dire dans la sélection de l'information qui est portée en légende. C'est pourquoi nous présentons dans les chapitres 7, 8 et 9 différentes familles de solutions qui visent d'abord à sélectionner l'information de flux qui sera cartographiée.

Dans le chapitre 7, la sélection des flux est appréhendée en fonction de variables spatiales liées d'une part, à la séparation des lieux sous une forme continue (la distance parcourue) et discrète (le voisinage) ; d'autre part, à leur proximité dans le cadre d'une relation d'appartenance à la même entité. La prise en compte de la composante spatiale a , de tous temps, nuit à la représentation des flux et des mouvements, en raison de la contrainte de localisation des lieux dans un espace géographique, nécessairement projeté et muni de deux dimensions. Ce problème spécifique à la représentation du mouvement explique le recours tardif au graphique pour les illustrer, il justifie aussi les stratégies d'évitement de la carte de flux déjà évoquées et place les solutions technologiques¹⁵⁶ comme famille centrale de solution. Nous allons montrer dans le chapitre 7 que l'espace peut être pris en compte sur une carte statique, par le biais de la distance parcourue par le flux.

Dans le chapitre 8, la sélection des flux consiste à appliquer un critère de sélection global (α) de manière à ne cartographier que les flux (F_{ij}) ou les mouvements (M_{ij}) qui répondent à ce critère que l'on considère *significatif*. Nous proposons pour cela une réflexion qui généralise la détermination de (α) en partant d'une suggestion de W. Tobler (1982) selon laquelle il est nécessaire d'observer la distribution des valeurs de flux (de la même manière qu'on observe la distribution de valeurs quantitatives relatives avant de les discrétiser), car l'allure de la courbe est supposée log-normale. Nous allons donc, dans un premier temps, chercher à déterminer la valeur de ce seuil, en adoptant une approche classique qui nous conduira à analyser $\log(F_{ij}^+)$ et $\log(M_{ij}^+)$, puis à les

¹⁵⁶ Les solutions d'ordre techniques et technologiques sont liées au développement d'un outil permettant de représenter le mouvement de manière active et non seulement statique.

transformer en fréquences cumulées. Nous proposerons ensuite de mobiliser un troisième graphique, une courbe de concentration des valeurs dans la mesure où, en réalité, (α) correspond à un double critère (ou à un unique critère qui agit doublement) qui agit sur la part de flux ou de mouvements (F_{ij}) ou (M_{ij}), en fonction de la part de figurés (L_{ij}) ; ce double critère global pouvant alors être estimé soit selon un objectif de *significativité de l'information* soit selon un objectif de *densité de la matrice*. Les solutions proposées conduisant à la suppression d'une partie des lieux de l'analyse, ceux dont les valeurs des liens (entrants ou sortants) ne sont pas suffisamment *significatifs*, il peut être souhaité de conserver tous les lieux. Pour cela, il convient d'appliquer un critère de sélection local.

Dans le chapitre 9, la sélection des flux est locale, c'est-à-dire focalisée sur l'analyse de la hiérarchie des lieux qui découle d'un système de flux dits dominants, elle mobilise les lignes et /ou les colonnes de la matrice (les lieux d'origine ou de destination des flux) en même temps que son cœur. L'approche mise en œuvre est fondée sur la méthode Nyusten-Dacey (1961), en raison de sa popularité mais aussi de la variété et de l'intérêt des solutions cartographiques de flux qu'elle propose. Construit en écho du précédent, ce dernier chapitre conduit à ne retenir qu'un et un seul flux (sortant ou entrant) par lieu dans l'objectif d'identifier *in fine* des régions nodales. Pour y parvenir, nous présentons, dans un premier temps, la méthode Nyusten-Dacey et les résultats (carto)graphiques qu'elle procure, et ceux qu'elle suggère. Nous procédons, dans un second temps, à une (re)construction cartographique de ces résultats qui nous conduit à examiner cinq étapes du processus de cartographie de *flux dominants*. Nous montrons ainsi que la méthode est, en fait, focalisée sur l'analyse des lieux à partir des valeurs de flux. Elle correspond à une *logique liens-lieux* plutôt qu'à la *logique de flux* classique. Le raisonnement diffère également : il est désormais qualitatif et non plus uniquement quantitatif, pour la représentation de la hiérarchie des lieux. Cela nous conduit à ouvrir une discussion sur les possibilités de sémiologie cartographique. Dans un troisième temps, nous proposons une validation empirique des solutions proposées et envisageons ses différents cas théoriques d'application liées aux possibilités de décomposition de la matrice présentées dans le chapitre 3. Si les auteurs ont mobilisé une matrice (F_{ij}), nous envisageons les conditions de l'application de la méthode Nyusten-Dacey aux matrices (F_{ji}) et (F_{ij}^+).

Les solutions proposées dans cette seconde partie se voulant générales et généralisables, elles seront présentées formellement à l'aide de la matrice de flux fictive servant d'exemple local (voir Figure 1-2, page 30). Elles seront ensuite appliquées sur des données empiriques : des matrices de flux variées s'exprimant à l'échelle locale et à l'échelle mondiale.

L'échelle locale sera représentée par la matrice des navettes intercommunales issues du fichier MOBPRO (INSEE, 2010) pour le département de l'Isère (38), dans l'ensemble des chapitres de cette seconde partie.

L'échelle mondiale sera, quant à elle, illustrée avec trois matrices de flux de personnes et de biens qui varieront en fonction des chapitres : une matrice de flux internationaux inter-états de marchandises (CEPII, 2007-2010) exprimés en valeur monétaire (chapitres 5, 6 et 8) une matrice de flux internationaux inter aéroports de passagers (OACI, 2001) dans le chapitre 7 et une matrice de flux commerciaux internationaux (OMC, 2007) dans le chapitre 9.

Toutes ces matrices seront présentées lors de leur première apparition dans le texte.

Enfin, ces solutions seront implémentées sous la forme de programmes développés sous RStudio© (intéressant spécifiquement les chapitres 3, 7, 8 et 9), disponibles en annexes.

5^{ème} CHAPITRE. Propositions de solutions liées aux aspects sémiologiques et esthétiques de la carte de flux : le renouvellement de la variable visuelle *Couleur*

« *Je suis contre la couleur qui camoufle l'incompétence et je reste contre tant que l'on croira qu'elle suffit pour représenter un ordre, qu'elle permet de superposer des caractères jusqu'à la limite de l'absurde, tant qu'elle coûtera inutilement des fortunes aux deniers publics dans une confusion entre connaissance et publicité, tant que, par son truchement, les responsables et la télévision diffuseront des images fausses et illisibles* »¹⁵⁷.

Jacques Bertin (1973).

Introduction du chapitre

Ce cinquième chapitre a pour objectif de présenter des solutions d'ordre graphique : esthétique et sémiologique, qui autorisent la cartographie d'une matrice de flux dense, naturellement sujette à *l'effet-spaghetti*. L'objectif est de représenter *toute l'information disponible* dans les différentes cases du cœur de la matrice, hors de la diagonale principale, sans opérer de sélection – ou d'élimination –, de réduction *a priori* des objets géographiques et/ou des valeurs à représenter.

L'approche que nous proposons intéresse, par conséquent, la question de la perception sensible des motifs des flux cartographiés, le caractère esthétique de la figure, c'est-à-dire de la carte de flux en elle-même et non comme un élément de la composition à laquelle elle appartient éventuellement. Si la perception des motifs principaux correspond à l'objectif de la carte, son caractère esthétique y contribue (l'ensemble définissant l'efficacité de la carte). Nous allons agir sur ce dernier, principalement à travers la mobilisation de deux variables rétinienne que sont la *Couleur* et la *Taille*. La première est appréhendée en tant que teinte sur le cercle chromatique et comme variable visuelle, tout comme la *Taille*.

Si la mise en œuvre de ces variables visuelles, appliquée au cas de flux, s'inscrit dans le cadre des principes formalisés par J. Bertin (1967), présentés dans le chapitre 4 (page 177), nous faisons l'hypothèse générale que **certains principes empiriques mis en œuvre dans le champ de l'informatique graphique, pour la visualisation de masses de données relationnelles sur un support écran, peuvent être transférés à la cartographie de flux**. Cette hypothèse conduit au bouleversement de l'esthétique de la carte de flux statistique et, plus loin, ouvre la voie vers **le renouvellement de la sémiologie graphique de la carte de flux** (Bahoken, 2014).

¹⁵⁷ Aujourd'hui, soit trente ans plus tard après cette citation, J. Bertin serait, bien évidemment, étonné de constater à quel point l'usage de la couleur n'est plus une limite, ni financière, ni technique pour l'édition électronique sur un support écran ; l'édition papier étant encore, parfois, délicate. L'utilisation de la couleur participe désormais, dans l'ensemble, de la définition de l'esthétique de la carte et aussi de son efficacité, que sa vocation soit à des fins publicitaires, de communication, d'édition voire même scientifique.

Pour valider cette hypothèse, nous mobilisons les méthodes de cartographie de flux présentées dans le chapitre 4, en fonction des familles de matrices définies dans le chapitre 3, en prenant notre petite matrice d'exemple comme cas d'illustration. Comme nous l'avons mentionné dans le chapitre 3, l'intérêt du recours à une décomposition de la matrice réside dans le caractère généralisable des méthodes de cartographie de flux applicables, en fonction des principales familles de matrices. En supposant que la famille de *matrices asymétriques*, correspond aux données observées (F_{ij}), elle décrit des flux bilatéraux qu'il est nécessaire de représenter à l'aide de doubles flèches, ces derniers étant nécessairement orientés. La famille de *matrices symétriques*, illustrée par (F_{ij}^+), correspond au volume bilatéral de flux, à représenter à l'aide de bandes non orientées. La famille de *matrices antisymétriques*, illustrée par (F_{ij}^-) décrit le solde bilatéral de flux ou le transfert net, soit par une seule carte où les figurés, par exemple (F_{ij}^-) >0 , pour éviter la redondance des figurés.

Les solutions liées aux aspects graphiques, esthétiques et sémiologiques examinées dans ce cinquième chapitre proposent de renouveler l'image cartographique des principales familles de matrices de flux présentées dans le chapitre 3. Elles s'inscrivent dans un cadre d'analyse restreint que nous limitons volontairement par les trois contraintes suivantes.

– La cartographie portera sur *l'ensemble de l'information disponible* dans la matrice, c'est-à-dire sur l'ensemble des flux par un nombre équivalent de figurés rectilinéaires pour les liens (ou ponctuels pour les lieux). Dans ce chapitre, les liens qui relient les couples d'origine et de destination seront illustrés par des traits droits. Le cheminement exact du flux entre les lieux n'étant pas connu, le figuré qui illustre ce flux n'a pas besoin d'être incurvé pour des raisons autres que liées à l'esthétique de la figure : nous ne lui appliquons aucune transformation¹⁵⁸ d'ordre graphique ou même cartographique. Ce trait sera, en revanche, assorti d'une pointe, lorsque le flux qu'il représente est orienté. Autrement dit, le trait sera remplacé par une flèche (ou tout autre procédé permettant de représenter une orientation) si l'orientation du flux a un sens. Pour représenter les flux observés du point de vue des lieux, nous faisons le choix du point. Le choix de ces formalismes simples se justifie par la facilité de leur mise en œuvre et leur automatisation aisée.

– La conception de la carte est nécessairement contrainte par la composante spatiale, sur laquelle nous n'intervenons pas. Cela signifie que la question du choix du système de projection ne sera pas abordée, pas plus que les transformations (carto)graphiques de la position et/ou de la morphologie des objets. Ces transformations qui agissent sur la position des lieux et/ou sur les liens correspondent à une famille de solutions que nous ne pourrions présenter ici. Nous supposons ainsi que le système de projection de la matrice dont il est question est, par conséquent, connu *a priori*. Il s'inscrit dans le repère euclidien traditionnel.

– Les données de flux observées au niveau des liens (i,j) et des N lieux étant de nature quantitative, les principes classiques de sémiologie graphique (Bertin, 1967) imposent le recours *a minima* à la variable visuelle *Taille*. **Appliquée aux flux, la Taille conduit d'une part, à une représentation de la largeur du figuré linéaire en proportion de la valeur (F_{ij}) qu'il représente et d'autre part, à son ajustement en fonction de la résolution de la carte, de manière à autoriser la perception simultanée du plus petit $\text{Min}(F_{ij})$ et du plus gros figuré $\text{Max}(F_{ij})$.** Cela nécessite donc de borner le tracé du plus gros figuré, qui correspond à la valeur maximum, de manière à éviter l'absorption visuelle des plus petits flux, et de conserver une harmonie de l'ensemble des figurés.

¹⁵⁸ Le lecteur intéressé par ces solutions qui conduisent à la transformation de l'apparence visuelle du lien pourra consulter, à titre d'exemple, les travaux réalisés par O. Bonin et al. (2014).

Nous introduisons toutefois la *Couleur* pour plusieurs raisons. Tout d'abord, parce qu'elle participe indéniablement de l'esthétique de la figure que nous présentons dans la première section. Nous transférons ensuite la variable visuelle *Transparence* depuis le champ de l'Informatique graphique à la cartographie de flux, de manière à pouvoir représenter l'ensemble de l'information disponible dans la matrice dont il est question. L'expérimentation est menée sur la matrice des volumes bilatéraux (Fij⁺), elle est illustrée par les navettes de l'Isère. La troisième section est consacrée à la mise en œuvre de la *Couleur*. Elle porte sur le transfert de la sémantique de l'opposition colorée au niveau des liens. La méthode est appliquée à la matrice (Fij⁻) et aux matrices complexes, elle est également illustrée par la matrice des soldes bilatéraux des navettes de l'Isère. La quatrième section permet de valider les solutions proposées, en montrant leur caractère généralisable. Pour cela, les suggestions sont appliquées à une matrice de flux commerciaux internationaux. Cette matrice *asymétrique* décrit l'ensemble des échanges commerciaux, en valeur (US \$), réalisés entre des pays ou groupes de pays dont la collection est issue de la base de données CHELEM du CEPII.

5.1 La nouvelle esthétique de la carte de flux

La carte de P. Mariot mentionnée par R. Brunet (voir Figure I-1-1, page 13) est réalisée en noir, sur un fond blanc. Elle correspond à l'esthétique classique ou usuelle de la carte de flux, replacée dans son époque conceptuelle et technologique, présentée dans son plus simple appareil. Aussi à une pratique classique replacée dans son époque, contextuelle et méthodologique.

L'esthétique est une notion universelle qui intéresse la question de l'apparence perçue et de la beauté. La carte ayant d'abord pour vocation d'informer sur un processus spatialisé, l'esthétique d'une carte intéresse aussi son efficacité dans la transmission du message. Si nous ne rentrons pas dans une réflexion approfondie sur l'esthétique en cartographie statistique, car elle a fait l'objet d'une thèse soutenue très récemment (Jégou, 2013a), nous pouvons en relever deux particularités qui permettent d'appuyer notre réflexion.

En matière de cartographie, l'esthétique est une notion abordée d'une part, au niveau de la géométrie des objets, qui se traduit sur la carte par la finesse, la précision des tracés et le degré de généralisation des contours ; de l'autre, dans le choix et l'harmonie des couleurs utilisées qui définissent son style général. L'esthétique peut aussi être appréhendée par la façon dont l'auteur de la carte parvient à transmettre un message complexe de façon simple, afin que la carte soit alors perçue comme efficace et « *agréable à regarder* », pour reprendre une expression de L. Jégou. D'après l'auteur, deux sources du caractère agréable associé à la lecture d'une carte peuvent ainsi être mentionnées : « [...] *le plaisir de la bonne compréhension et de l'acquisition d'informations (en état conscient) et, de l'autre, celui de la perception de l'harmonie d'une image, dans ses composantes visuelles donc en grande part l'harmonie de ses couleurs (en partie inconscient). Le premier plaisir est fortement dépendant de la lisibilité et se voit renforcé par le second* »¹⁵⁹.

Les résultats de l'enquête *Cartographie et esthétique* menée par L. Jégou en 2013, auprès de 1 400 internautes, sont, en ce sens, éloquentes : pour 80% des enquêtés, « *une carte visuellement simple est efficace* » (Jégou, 2013b : 41), mais seuls 49% des répondants considèrent que les aspects de sémiologie jouent un rôle dans le fait qu'une carte soit agréable à regarder. Nous supposons que cela signifie que cet aspect est rempli par un autre critère qui est, en l'occurrence, l'utilisation de la couleur. Ainsi, le bleu clair Le côté agréable d'une carte apparaît, en effet, fondamental pour 85% des enquêtés, il s'agit d'un aspect « [...] *important [de la carte] pour son efficacité* » (Jégou, 2013b : *ibid*). Cependant, l'efficacité de la carte varie en fonction de la perception de son côté « *agréable à regarder* » et, par conséquent, de sa teinte. Dans la sémantique d'opposition colorée, par exemple, si

¹⁵⁹ Ces précisions nous ont été apportées dans le cadre d'une conversation informelle avec l'auteur, en 2015.

l'utilisation classique du bleu est approuvée par 80% des enquêtés pour représenter des valeurs négatives, 55% d'entre eux considèrent cette teinte froide comme agréable. Pour les valeurs négatives, le bleu clair est considéré comme plus efficace que le vert clair par 72% des enquêtés, mais, en même temps, le vert clair est ressenti comme plus agréable (55% contre 18% pour le bleu clair). « *Cela explique l'idée selon laquelle le caractère agréable est ici dissocié de l'efficacité de la carte, dans ce contexte précis de la couleur froide associée à une variation négative* »¹⁶⁰. Elle conduit l'auteur à la conclusion selon laquelle « [...] *l'aspect agréable de la carte ne participe pas à son efficacité perçue* » (Jégou, 2013b : 41).

L'aspect esthétique d'une carte résulte en pratique de choix qui sont réalisés par le cartographe au niveau graphique, tout au long du processus de réalisation de la carte. Ces choix intrinsèquement liés à la sensibilité de l'auteur sont, par conséquent, d'ordre cognitif ou intellectuel ; ils sont aussi d'ordre sémiologique car ils s'inscrivent dans le cadre d'une formalisation graphique aujourd'hui bien établie et sur laquelle nous ne reviendrons pas (Bertin, 1967 ; Tufte, 1986). Ainsi, les choix en matière d'esthétique concernent, en particulier, trois aspects principaux du processus de production : ils intéressent les primitives graphiques, c'est-à-dire le dessin et le niveau de précision des tracés ; la palette et la gamme chromatique utilisée pour la représentation des figurés, éventuellement leur signification ; enfin, le style général de la figure, conférant à la carte une certaine originalité (Christophe, 2012a, 2012b).

Notre recours à l'esthétique n'est toutefois pas gratuit. Nous verrons dans les deux sections suivantes comment le renouvellement de l'esthétique que nous présentons, dans cette première section, lorsqu'elle est mobilisée en lien avec les principes de sémiologie cartographique classiques (Bertin, 1967) conduira d'une part, à bouleverser les codes de la carte de flux usuels et d'autre part, à renouveler de manière saisissante sa sémiologie graphique. Nous tenons à rappeler que notre propos intéresse la carte de flux en elle-même et non comme un élément d'une composition plus large, comportant éventuellement d'autres éléments (du texte, des graphiques, ...).

5.1.1 *Le choix de la couleur*

La cartographie des données disponibles dans n'importe laquelle des matrices de flux (voir chapitre 3) nécessite de placer sur une carte deux familles de figurés : la première, pour illustrer les flux observés entre les lieux et la seconde, pour représenter l'effet des flux sur la distribution des lieux. La méthode de cartographie la plus simple à mettre en œuvre est celle qui décrit le volume de flux : elle mobilise des bandes et points qui sont proportionnelles aux valeurs correspondantes dans la matrice. Cette méthode est celle des matrices symétriques, telles que (Fij⁺).

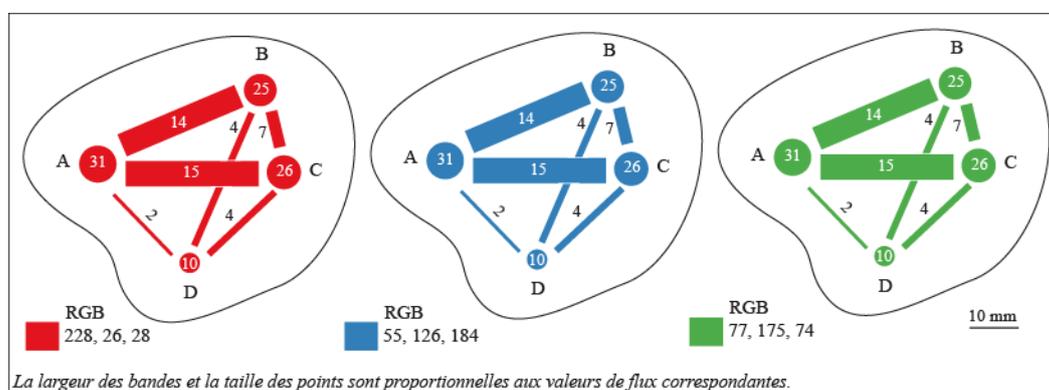
Sur le plan sémiologique, la cartographie de (Fij⁺) fait appel à une série statistique positive. Elle ne nécessite donc pas l'utilisation d'une couleur particulière, au sens de teinte. C'est aussi la raison pour laquelle l'essentiel des cartes est réalisé en noir-et-blanc : la sémiologie classique de la carte de flux ne nécessite pas l'utilisation de la variable visuelle *Couleur*. Celle-ci étant traditionnellement réservée à la représentation de variables continues discrétisées en *k* plages de valeurs, souvent assortie à la *Valeur*. Nous en concluons que le recours à la couleur est inutile, lorsqu'elle n'est pas mobilisée avec une autre variable visuelle (ou une autre dimension statistique), ce qui est le cas spécifique de la cartographie du volume bilatéral de flux.

La carte du volume bilatéral (voir Figure 3-15, page 130) est réalisée dans un style en noir-et-blanc. Elle présente l'aspect classique des figures anciennes réalisées à main levée. Son efficacité visuelle réside dans l'équilibre des deux teintes (noire et blanche), mais pas seulement. Si le Noir est réservé aux figurés, il est aussi associé au Blanc, sa teinte opposée, dans l'objectif de renforcer la propriété de

¹⁶⁰ Les précisions suivantes nous ont également été apportées par l'auteur, en 2015.

séparation de la couleur. Ce choix du noir-et-blanc est purement subjectif. En effet, si nous prenons l'exemple du volume bilatéral de flux, nous constatons que la modification de la teinte, utilisée en opposition au Blanc, est tout aussi lisible (voir Figure 5-1).

Figure 5-1. Illustration du caractère facultatif de la teinte, pour cartographier le volume de flux (F_{ij}^+)



La Figure 5-1 illustre l'invariance de la perception des motifs de flux cartographiés, lorsque les figurés sont colorés et le fond porté en blanc. Nous en concluons que le choix subjectif de la teinte, pour cartographier une matrice de flux symétrique, est un critère négligeable, au regard de l'importance de la *Taille*¹⁶¹ ; l'introduction d'une teinte colorée serait davantage liée à l'attrance qu'elle suscite. Nous validons ici l'un des résultats de l'enquête réalisée par L. Jégou (2013) selon lequel le choix de la teinte qui conditionne « [...] *l'aspect agréable de la carte ne participe pas à son efficacité perçue* » (Jégou, 2013 : 41), mais y contribue. En effet, L. Jégou nous précise¹⁶² que le fait d'utiliser une couleur plutôt que le noir-et-blanc dans une carte de flux qui n'en a pas besoin permet potentiellement de jouer sur le côté esthétique et perceptif général, de dégager la carte du fond d'un document plus grand (ou en plaçant une rime graphique issue d'une charte plus générale, etc.). Dans ce cas précis, la couleur joue alors un rôle d'identification plutôt que d'information, mais pas à l'intérieur même de la carte, ce qui nous conduit à sortir de notre cadre d'analyse. De ce fait, pour reprendre une expression courante de C. Zanin, ce qui « *saute aux yeux* » dans une carte classique du volume de flux, c'est bien la perception des différences plus ou moins importantes de la taille des cercles et des bandes, c'est-à-dire un critère d'ordre méthodologique lié à la sémiologie des flux plutôt qu'à leur esthétique. Cela tient aussi au fait que l'œil perçoit, avant toute chose, des rapports de proportionnalité.

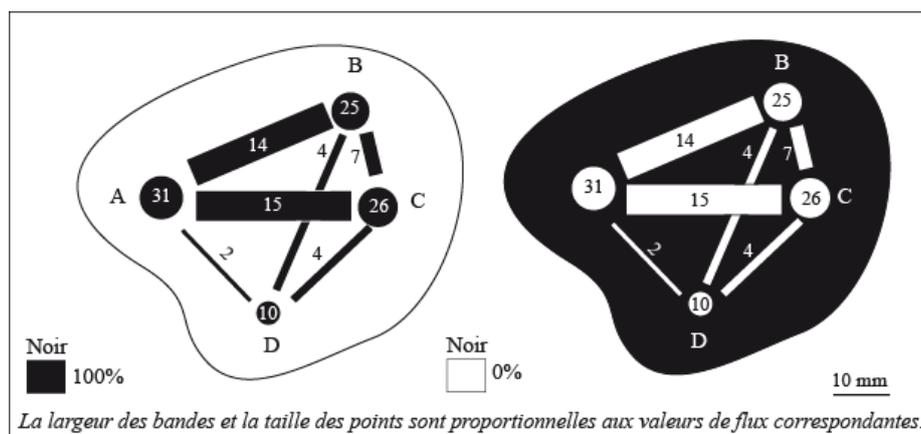
Le fait de considérer que l'utilisation de la couleur n'apporte pas d'information particulière à la perception des motifs d'une carte du volume de flux nous conduit à proposer un changement de style : en mobilisant toujours les teintes noires et blanches, en délaissant momentanément la couleur. Le Noir et le Blanc étant deux teintes opposées et complémentaires, il est possible d'invertir leurs positions, de substituer l'une à l'autre sans nuire à la transmission du message. Seule l'esthétique de la figure est modifiée puisqu'elle passe du noir-et-blanc des figurés noirs sur un fond blanc, au blanc-et-noir des figurés blancs sur un fond noir.

La Figure 5-2 compare la figure classique du volume de flux (à gauche) à son négatif (à droite). Le style classique du noir-et-blanc est remplacé par celui du blanc-et-noir, tous deux correspondant en pratique à une palette en niveaux de gris.

¹⁶¹ Nous reviendrons sur l'importance de cette variable visuelle dans le chapitre 6.

¹⁶² Cette précision nous est apportée lors d'une conversation informelle.

Figure 5-2. Vers une nouvelle esthétique de la carte de flux



Il s'agit là d'un premier bouleversement de la figure classique du volume de flux qui résulte de l'inversion de la teinte du fond de la carte et, par conséquent, de celle des figurés. Plus généralement, cette inversion conduit à faire le choix d'une teinte sombre pour le fond de la carte et d'une teinte claire pour les figurés.

L'intérêt de ce choix esthétique illustré par la Figure 5-2 est double : il permet d'une part, de laisser les figurés en blanc, de les réserver à l'utilisation d'autres variables visuelles telles que la *Valeur*, par exemple, pour une coloration discrétisée des figurés ; de l'autre, d'améliorer la perception des motifs : G.Tabart et al. (2007) ayant pu démontrer que les contrastes étaient mieux perçus, lorsque le fond est noir.

Pour convaincre de l'intérêt de ce passage en négatif du fond de la carte, en lien avec l'ouverture des possibilités de renouvellement d'ordre sémiologique que nous montrerons, nous présentons dans la section suivante l'intérêt de l'utilisation d'un fond sombre.

5.1.2 De l'intérêt d'un fond sombre

Cette première sous-section intéresse le niveau de perception global de la carte de flux, focalisé sur le fond de la planche, sur le style ou l'apparence de la carte que nous proposons de teinter de *Noir*. En cartographie statistique, la teinte du fond de carte est généralement claire, voire blanche. Ce choix n'est pas anodin, il s'explique par des raisons historiques d'ordre théorique et pratique qui sont évidentes, nous les rappelons toutefois pour mémoire.

La couleur blanche était historiquement privilégiée pour figurer en fond de page des cartes statistiques dites thématiques. Elle est héritée d'une cartographie réalisée d'abord manuellement, puis majoritairement en noir et blanc, surtout proposée sur un support papier, les possibilités manuelles d'usage de la couleur étant *a priori* moins étendues que sur un support écran. Sur le plan théorique, le choix d'un fond – de page ou de carte – blanc voire clair permet de représenter avec une teinte opposée, généralement le noir, les figurés ainsi que les différents éléments qui composent la planche cartographique (les titres et sous-titres, l'échelle, ...). Sur le plan pratique et pour des raisons économiques liées au prix de l'encre, *a fortiori* de la couleur, il était préférable de réserver l'usage des teintes blanches aux surfaces les plus importantes de la planche qui ne nécessitaient pas d'encre et d'utiliser le noir – voire les couleurs – à la représentation des figurés, hachures et autres poncifs. Aujourd'hui, la situation est quelque peu différente puisque l'impression en couleurs s'est démocratisée. Le recours aux teintes colorées demeure toutefois plus onéreux que l'utilisation de la couleur noire, elle n'est pas toujours stable à l'impression. Dans bien des cas, l'utilisation de la couleur, en cartographie statistique n'est pas justifiée autrement que par des raisons esthétiques et

d'attraits. C'est pourquoi nous faisons le choix, dans un premier temps, de recourir à une palette en noir et blanc. Cela va nous permettre de démontrer l'intérêt de l'utilisation du *Noir* (et blanc) et le caractère esthétique de ce nouveau style de carte de flux, puis d'introduire dans le champ de la cartographie statistique, une nouvelle variable visuelle telle qu'elle est suggérée par A. Mac Eachren (1995), la *Transparence*.

5.1.2.1 *Éléments sur l'esthétique du Noir*

Le *Noir* nous est apparu d'une part, comme l'une des teintes permettant d'augmenter la lisibilité globale des cartes de flux de manière sensible – au sens où elle favorise leur perception cognitive – et de façon significative, car elle permet de renforcer la *saillance* des motifs. D'autre part, comme un moyen de réaliser à peu de frais des cartes de flux que nous jugeons esthétiques et, comme tout jugement, ce choix est purement subjectif. Nous montrerons que l'inversion de la teinte du fond de la carte, que nous allons opérer, n'a pas pour conséquence de modifier fondamentalement la perception des motifs de la carte de flux.

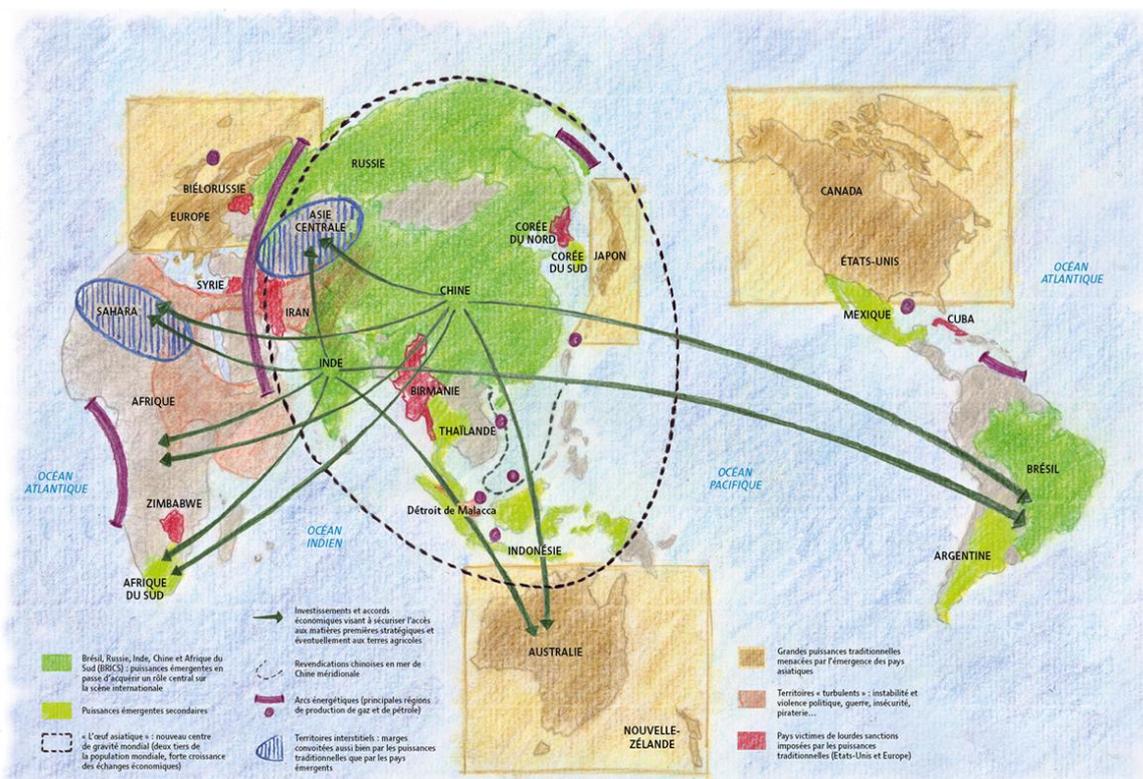
L'esthétique est une notion *mouvante* dans le temps et dans l'espace. Chacun sait que, pour un même territoire, les canons de beauté d'hier ne sont pas ceux d'aujourd'hui, certains d'entre eux ne sauraient d'ailleurs être comparés ne serait-ce qu'en raison des outils mobilisés pour les représenter. Cela ne signifie pas que nous pensons que l'outil informatique ait engendré une révolution esthétique puisqu'il n'intéresse qu'une partie de la création artistique. De notre point de vue, l'outil informatique – en permettant de gérer de manière automatisée des données manipulées auparavant manuellement – ne fait que modifier notre perception des relations colorées et notre niveau d'attention sensible, en même temps qu'il décuple certaines de nos possibilités graphiques, pour peu que l'on soit familier avec les codes liés à son utilisation. C'est en ce sens que le défi de P. Rekacewicz, journaliste politique se réclamant radical et cartographe au sein du Monde Diplomatique, de réaliser l'ensemble de ses cartes à la main, sans recourir à l'outil, est réussi. La posture assumée¹⁶³ de l'auteur constitue d'ailleurs un véritable pied de nez au *tournant géo numérique* et à la pratique bien répandue en cartographie dite du *presse-bouton*, laissant le choix aux développeurs, souvent informaticiens, de décider du style à appliquer aux cartes et des méthodes informatiques pour les mettre en œuvre et ce, en faisant fi des principes universels de la perception visuelle.

La Figure 5-3 présente une esquisse du dossier *Cartes en colère*¹⁶⁴ réalisé par P. Rekacewicz pour l'Atlas 2012 Mondes Emergents, du *Monde Diplomatique*. Elle décrit « l'œuf asiatique » qui caractérise la croissance exceptionnelle de cette région du monde. D'après l'auteur, « *cette formule illustre bien les tendances actuelles de l'économie et du commerce mondiaux. La Chine « s'interconnecte » très étroitement avec des pays qui furent et restent parfois de grands rivaux, comme les Etats-Unis ou le Japon* ».

¹⁶³ Nous faisons allusion ici à une conversation informelle avec l'auteur sur la pratique actuelle de la cartographie statistique.

¹⁶⁴ Ce dossier est accessible sur le blog Visions Carto de l'auteur, à l'adresse suivante : <http://blog.mondediplo.net/2013-12-22-Cartes-en-colere>

Figure 5-3. 'L'œuf asiatique' (Rekacewicz, 2012)



Source : *Monde Emergents*, 2012.

Il est aussi de notoriété publique que les critères de beauté varient en fonction des aires dites culturelles. Ils dépendent donc de la localisation géographique du cartographe, de l'observateur, de leurs sensibilités personnelles mais aussi de la vision collective de la société à laquelle ils appartiennent. M. Pastoreau (2008) a d'ailleurs bien montré dans son ouvrage consacré au *Noir*, l'évolution de la manière dont cette teinte est considérée en fonction des périodes, tantôt comme une couleur, tantôt non : tout dépend « [...] d'enjeux économiques, politiques, sociaux, symboliques s'inscrivant dans un contexte précis [...] » (Pastoreau, 2008 : 17). Sans rentrer dans les détails d'un débat fort intéressant et minutieusement documenté par l'auteur, il est intéressant de noter que l'ouvrage apparaît, dans l'histoire sociale de cette couleur, au moment même où son usage se singularise dans le champ de la visualisation graphique des données. En matière de cartographie, la tendance actuelle est, en effet, celle de la nuit. Elle se manifeste par l'association d'un fond sombre, souvent noir, plutôt mat que brillant, présenté avec des figurés plutôt brillants voire lumineux. L'image publicitaire de l'entreprise Google®, pour promouvoir la nouvelle version de son moteur de recherche, est en ce sens symptomatique de cette tendance (voir Figure 5-4).

Figure 5-4. L'esthétique de la couleur noire



Source : T. Joliveau, 2013.

D'après Google©, l'utilisation d'un fond noir ne serait pas seulement esthétique : elle contribuerait aussi à réaliser une économie d'énergie de 20% au niveau des serveurs de stockage par rapport au fond blanc, d'où une incitation dans le cadre d'une campagne publicitaire dédiée à l'utilisation de *Google Black Out*¹⁶⁵. A noter que l'utilisation du même fond noir sur un support numérique destiné à l'impression papier produit l'effet inverse : une dépense considérable d'énergie.

Dans la culture populaire, la couleur noire n'est pas associée à la protection de l'environnement comme Google© veut bien nous le faire croire. Elle est plutôt classiquement associée à la pollution : le *smog* londonien ainsi que les marées noires représentent des atteintes à l'environnement reconnues, par tous, comme significatives. Si dans la culture populaire le *Noir* revêt différentes acceptions symboliques, l'une¹⁶⁶ d'entre elles fait référence à l'élégance, à la sobriété, au luxe ainsi qu'à la *classe*, celle-ci renvoyant à la fois à une apparence distinguée relative à une catégorie sociale, à une culture : celle des *riches* qui peuvent s'offrir les objets de luxe associés à leur rang social ; les *bobos*, souvent des *riches* qui se cachent, multiplient d'ailleurs l'utilisation combinée de couleurs de l'arc-en-ciel, marquant ainsi leur appartenance à une classe sociale spécifique et distincte d'une autre catégorie de riches qualifiée péjorativement de *bling-bling* qui, elle, n'a pas d'états d'âme pour recourir aux vêtements sombres. Les publicitaires ne s'y sont pas trompés : l'emballage de ces objets de luxe est souvent noir, comme dans le cas des parfums Chanel©, à moins que ce ne soit l'objet lui-même. Si le caviar en est l'un des illustres représentants, on peut aussi mentionner les stylos Mont Blanc© et l'i-phone de Apple© comme accessoires tendance, le *smoking* et la *petite robe noire* de Chanel© comme vêtements d'apparat historiques qui participent tous de la symbolique positive du *Noir*.

Ces différents objets luxueux de couleur noire participent d'une esthétique d'ensemble que la culture populaire et les médias en *vogue* – sans jeu de mot – qui leur sont associés considèrent comme l'excellence, le beau. Ils correspondent aussi à des objets d'usage nocturne, la couleur noire étant par ailleurs caractéristique de la nuit qu'il convient d'éclairer pour les voir, mieux les mettre en valeur. C'est pourquoi les objets brillants sont souvent associés à ceux qui sont noirs (les parures en diamants des *stars* à la petite, ou plutôt longue, *robe noire*), à la nuit et pour cause : la lumière artificielle est leur meilleure alliée. On voit d'ailleurs bien, sur la Figure 5-4, l'intérêt du fond noir pour renforcer la

¹⁶⁵ Le moteur de recherche en format *Black Out* est accessible aux adresses suivantes : <http://www.gllgle.fr/> ou <http://www.blackl.com/black-google.php> [liens vérifiés le 02/08/2015].

¹⁶⁶ La couleur noire ayant aussi son versant sombre, diabolique, fasciste, raciste ou encore maléfique, nous ne tenons pas à l'envisager ici pour des raisons symboliques, mais aussi politiques et cela, indépendamment du fait qu'elle ait pu être mise en œuvre pour réaliser des cartes de flux par un certain parti politique dont nous préférons taire le nom.

représentation des espaces qui ont accès au (ou qui utilisent massivement le) moteur de recherche Google© : ils sont représentés par une teinte jaune brillante, luminescente qui illustre l'éclairage dont les espaces concernés font l'objet. Ces zones *éclairées* sont, en réalité, *illuminées* par un *astre* tout puissant : la Figure 5-4 montrant, avec insistance, l'utilisation symbolique de la lumière apportée par le soleil, cette lumière qui fait référence à la *connaissance* à laquelle ont accès certains, et non ceux qui sont justement dans le noir, telle un message subliminal de l'allégorie de la caverne (Platon : République VII). Nous notons aussi la représentation du soleil, en pleine nuit, sous une forme particulièrement brillante, à la limite de l'éblouissement et qui n'éclaire qu'une partie de l'humanité. Cette utilisation particulière du *Noir* n'est pas l'apanage des publicitaires, elle est aussi mise en œuvre par des cartographes, qu'ils soient professionnels ou en herbe, ce qui illustre la manière dont les communicants déteignent sur les cartographes.

5.1.2.2 La communication qui déteint sur la cartographie des réseaux

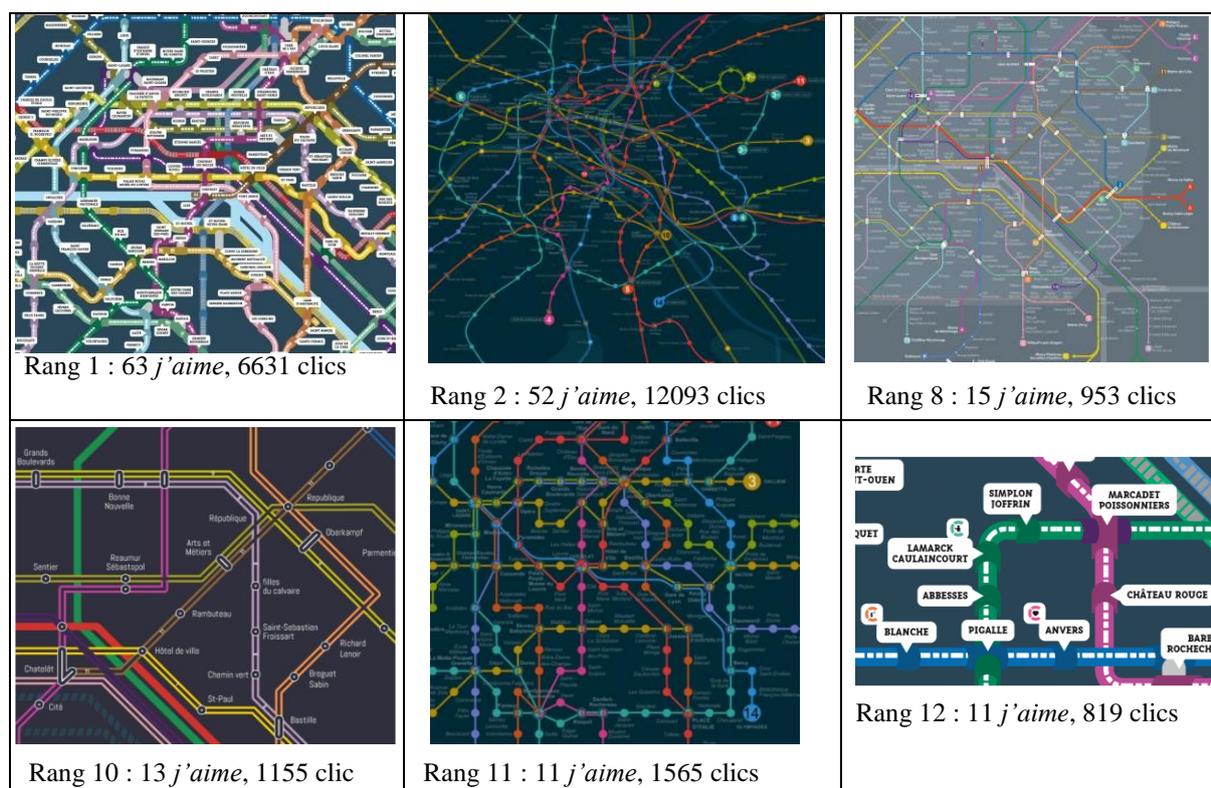
La couleur noire, telle que mise en œuvre par les publicitaires, est de plus en plus utilisée par les cartographes. Il faut dire que l'alliance couleur noire / couleur brillante est souvent heureuse, d'où le gage d'esthétique des figures qui la mettent en œuvre, qu'elles soient le fait de professionnels ou de graphistes en devenir, mais aussi du grand public, pour le volet perception de l'esthétique donnée à voir.

Un bon exemple de cette affirmation est apporté par les résultats à *Check my Map*¹⁶⁷, *Concours de création de la carte de métro de Paris* organisé en 2011 par *CheckMyMetro* et *Creads*, une agence militant pour l'ouverture des données publiques (l'*OpenData*) en France. Les plans proposés semblent avoir été jugés par des internautes, en fonction de leur caractère esthétique mais aussi de leur efficacité dans la transmission de l'information de repérage au sein de ce réseau ferroviaire. Ils sont supposés être aisément lisibles sur un *smart phone*, et se distinguer significativement du plan en vigueur réalisé par l'agence gestionnaire du métro parisien (la Régie Autonome des Transports Parisiens).

Il est intéressant de noter que le nombre maximal de votes en ligne pour les figures de ce concours, exprimé par la quantité de *clics*, est attribué aux versions du plan qui présentent un fond sombre voire noir, tout comme le nombre le plus important de *j'aime* qui est enregistré sur la page *Facebook*© correspondante. Les deux premiers plans exhibent d'ailleurs un fond bleu pétrole. La Figure 5-5 présente les six premiers plans de métro représentés sur un fond sombre. Sur les 20 premiers plans de ce concours, en termes de nombre de *clics*, qui présentent aussi le plus d'engagements de type *j'aime*, 9 d'entre eux, soit près de la moitié, sont réalisés sur un fond sombre et mat. C'est dire si le Noir est symptomatique d'une tendance générale.

¹⁶⁷ La liste des 83 cartes en compétition, ainsi que le nombre de votes dont elles font l'objet, est accessible à l'adresse suivante : <http://www.checkmymap.fr/creation-design/carte-metro-paris> [consultée le 11/10/2013 et vérifiée le 02/08/2015]. Signalons que certaines cartes sur fond sombre n'ont fait l'objet d'aucun engagement de type *j'aime* ; La raison tient à leur caractère farfelu, même si elles sont esthétiques ; les internautes ayant visiblement favorisé les figures sobres et efficaces qui permettent une lecture instantanée à l'écran, certaines d'entre elles ayant par ailleurs respecté la proximité usuelle des stations de métro (cartes de rang 1 et 2, notamment).

Figure 5-5. Sélection de cartes du réseau de métro parisien proposées sur un fond sombre, dans le cadre du concours de cartographie *Check my Map*©



Remarque : Le nombre de *j'aime* et de clics est relevé le 11 octobre 2013.

Source : Détails d'une sélection de cartes disponibles sur le site <http://www.checkmymap.fr/>

Etant donné qu'« *une couleur ne vient jamais seule* » (Pastoureau, 2008 : 17), le fond noir des cartes actuelles est souvent associé aux couleurs lumineuses ou brillantes, l'ensemble étant supposé plus esthétique que l'association couleur noire / couleur mate ou encore couleur blanche / couleur brillante. Les teintes colorées mates et pâles étant, quant à elles, considérées en Occident comme mieux adaptées à la journée, à la clarté du fond souvent représentée par la couleur blanche. Pour s'en rendre compte, nous proposons de discuter l'apparence visuelle d'une carte de flux selon que le fond du document est porté en noir ou en blanc, au regard de considérations esthétiques telles que nous les avons définies (beauté, perception et motifs).

L'esthétique qu'on sait être d'abord subjective, mais aussi collective, est soumise à des règles lorsqu'elle est envisagée dans le domaine de la production artistique. Soumise aux sens, l'esthétique caractérise ce qui est *beau*, agréable à voir. Elle intéresse le point de vue du cartographe, sa sensibilité et sa connaissance des règles de sémiologie graphique qui se traduit par des choix liés à son acception de la beauté, à une certaine utilisation de la couleur pour aboutir à un style particulier. L'esthétique intéresse aussi le point de vue de l'observateur lié à sa perception, et au jugement de la beauté donnée à voir par le cartographe ; les deux visions n'étant, bien entendu, pas toujours concordantes. Dans le domaine artistique, les choix réalisés en matière d'esthétique sont nécessairement effectués au regard de règles générales plus ou moins formalisées. Certaines sont liées à une acception du bon sens, souvent culturelle¹⁶⁸, liée à l'attrait pour certaines couleurs ; d'autres ont un objectif affiché d'efficacité dans la transmission du message, et font référence à une vision plus scientifique de l'usage

¹⁶⁸ Dans le domaine de la mode, par exemple, c'est une hérésie que de porter un costume noir avec des chaussettes blanches réservées à l'apparence sportive.

de la couleur, en particulier pour l'extraction et la transmission d'une sémantique d'opposition ou d'harmonie liée au mélange de couleurs. Ce mélange n'est en effet pas neutre, l'utilisation de couleurs complémentaires étant, par exemple, réputées se renforcer : c'est le cas de la synthèse de couleurs soustractives (en Cyan, Magenta, Jaune et Noir) où les couleurs opposées (teintes froides / teintes chaudes) sont opposées sur le cercle chromatique mais complémentaires : c'est le cas du vert / rouge ou encore du bleu / orange.

Pour comprendre la manière dont le recours au fond *Noir* peut renforcer la saillance des figurés clairs d'une carte du volume des flux, nous mobilisons une propriété de la teinte qui n'est pas mise en œuvre, à notre connaissance, en cartographie statistique des flux. Il s'agit de la transparence de la teinte, ou de son opacité.

5.2 L'introduction de la variable visuelle *Transparence*, l'exemple du (Fij⁺)

Dans la sémiologie graphique de la carte statistique, la variable rétinienne *Couleur* est mobilisée en tant que teinte sur le cercle chromatique et comme variable visuelle, principalement en association avec la *Valeur*. C. Zanin nous signale sa préférence pour le terme de *Valeur* qui renvoie inévitablement à un dégradé de teinte ou de ton, en différents paliers, plutôt qu'envisager cette double acception de la couleur que nous proposons. Ces deux utilisations de la couleur sont examinées, pour la première (la teinte), dans la présente section et pour la seconde, en tant que variable visuelle associée à la *Valeur*, dans le chapitre 6.

Concernant l'utilisation de la couleur en tant que teinte, nous conservons la suggestion du blanc pour les figurés, étant donné qu'ils seront disposés sur un fond sombre. Cela nous permet aussi de libérer un espace pouvant potentiellement accueillir une autre variable visuelle : une texture ou un *Grain* comme cela se fait couramment, ou bien une nouvelle variable visuelle, en l'occurrence la *Transparence*.

D'après le synopsis des variables visuelles formalisées par J. Bertin (1967), la *Couleur* (teinte) est une variable de séparation de l'image tout comme la *Taille*. Son utilisation classique n'introduit toutefois pas de modification morphologique de l'objet sur lequel elle s'applique, à l'instar de la *Taille*. La *Couleur* ayant une propriété associative, elle permet d'introduire une séparation, c'est-à-dire une distance visuelle entre les différents objets colorés, sans pourtant en modifier la disposition ni la forme, comme le fait la *Taille*.

Notre objectif dans ce cinquième chapitre étant de réaliser une carte de flux qui décrive *l'ensemble de l'information disponible* dans la matrice, celle-ci doit autoriser la perception de tous les flux : des gros figurés qui correspondent aux flux présentant des valeurs élevées et de la masse de petits figurés qui caractérise les flux présentant des valeurs faibles, ces petits flux étant pratiquement toujours éliminés de la représentation, ce que nous souhaitons éviter. Pour y parvenir en utilisant des solutions exclusivement d'ordre sémiologique, nous proposons d'agir sur les propriétés de séparation de la *Couleur*, grâce à l'outil informatique. Nous faisons l'hypothèse que **les principes empiriques mis en œuvre dans le champ de l'Informatique graphique, pour visualiser les masses de données relationnelles, sont susceptibles d'enrichir et de renouveler la sémiologie de la carte de flux.**

Plusieurs auteurs ont, en effet, proposé des extensions des principes de sémiologie graphique classiques de J. Bertin (1967) qui permettent de renforcer la perception des motifs, en lien avec les nouvelles possibilités de visualisation offertes par le support-écran. Pour ce qui est de la carte de flux, deux possibilités d'action au niveau de la teinte sont fructueuses. La première consiste à appliquer la variable visuelle *Transparence* suggérée par A. Mac Eachren (1995), elle fait l'objet des sous-sections suivantes. La seconde possibilité vise à transférer à la cartographie de flux, la sémantique de l'opposition colorée telle qu'elle est mise en œuvre en cartographie statistique, pour la représentation de stocks : elle est présentée dans la section 5.3.

Les propositions seront toutes appliquées à la matrice de l'Isère, que nous introduisons dans la sous-section suivante, avant de les présenter.

5.2.1 Présentation des données

La matrice de l'Isère est issue du fichier MOBPRO de l'INSEE 2010, proposant un ensemble de variables sur un échantillon d'individus actifs de 15 ans ou plus ayant un emploi, se déplaçant quotidiennement entre leur lieu de résidence et leur lieu de travail. Chaque enregistrement correspond à un individu décrit selon ses principales caractéristiques sociodémographiques, celles du ménage auquel il appartient et celles de son déplacement. La localisation des lieux de résidence et travail est la commune, hormis les communes de Paris, de Lyon et Marseille renseignées au niveau de l'arrondissement municipal. Seuls les actifs ayant un emploi et résidant dans une commune française sont recensés. Ces emplois sont à temps complet ou à temps partiel, rémunérés ou non, ils incluent aussi les individus ayant une « activité réduite », les chômeurs, les étudiants, retraités salariés ainsi que les militaires du contingent.

Les incertitudes qui caractérisent ce fichier sont présentées dans l'Encart 5-1

Encart 5-1. Le fichier de détail MOBPRO de l'INSEE

Plusieurs éléments caractérisent le fichier MOBPRO. En effet, si le lieu de travail concerne communément la commune de localisation de ce dernier, ce n'est pas le cas pour certains professionnels mobiles, tels que les chauffeurs routiers : leur lieu de travail correspond à leur lieu de résidence. Une seconde particularité est liée au mode déclaratif du lieu de travail susceptible d'engendrer des erreurs et à l'orthographe, parfois illisible voire incomplète ou contradictoire avec le code INSEE déclaré. Cette incertitude du lieu de travail est également liée à la taille de la commune, en raison d'une variation de la précision du RP. Ainsi, d'après l'INSEE : « Pour les communes de moins de 10 000 habitants, l'exploitation principale fournit des résultats plus précis car elle porte sur l'ensemble des bulletins collectés. Pour les communes de 10 000 habitants ou plus, la précision est à peu près équivalente pour les deux Exploitations ». Enfin, il convient de noter la pondération apportée à la valeur du flux (Fij, par la variable (IPONDI) fournie avec jusqu'à 15 décimales, pour assurer la représentativité de l'information. L'INSEE fournit le détail des opérations réalisées sur les fichiers issus du recensement mis à disposition. Elles concernent : 1 - Les contrôles de la collecte ; 2 - La saisie ; 3 - L'organisation des traitements ; 4 - La codification ; 5 - Les redressements automatiques ; 6 - La recodification et la validation des résultats qui donne lieu à la création des variables qui sont diffusées.

La matrice que nous mobilisons est une extraction du fichier MOBPRO pour l'Isère. Le département comptant 533 communes, cette matrice est par essence fortement soumise à *l'effet spaghetti* puisqu'elle compte (533*532) cases hors diagonale, soit 283 556 valeurs théoriques. Cette matrice pourtant fermée (elle est limitée aux flux échangés entre les communes du département, voir chapitre 1) n'étant pas « carrée » : les flux s'expriment entre (532*498) communes. C'est pourquoi nous l'avons transformée de sorte que le nombre de lignes et de colonnes soit identique. L'Annexe R-1 décrit cette procédure et présente une carte de l'information théoriquement disponible. L'image qui en résulte n'ayant pas d'intérêt, nous n'allons pas la présenter. Nous allons plutôt décomposer cette matrice et ne retenir, dans un premier temps, que la composante (Fij⁺), également illisible. C'est pour remédier à cette complexité graphique que nous allons agir sur la teinte, en l'occurrence sur le canal Alpha que nous présentons dans la sous-section suivante.

5.2.2 Le canal Alpha de la teinte

Une manière intéressante d'améliorer sensiblement la perception de masses de figurés visualisés sur un support écran consiste à agir sur leur teinte, en modifiant ses propriétés. La teinte étant toujours définie dans un espace colorimétrique particulier, nous prenons l'exemple du mode Rouge, Vert, Bleu (R, V, B) qui correspond au mode image/écran, à la différence du mode MSN orienté édition papier.

En cartographie de flux, seules les propriétés de saturation, de teinte et de luminosité du (R, V, B) sont traditionnellement mobilisées pour définir l'apparence colorée d'une carte. Or, depuis la généralisation du support écran, il est désormais possible d'agir sur le canal alpha (R, V, B, A) d'une nuance de teinte, de manière à ajouter un effet de *Transparence* ou d'opacité à son rendu.

Le canal alpha conduit, en effet, à une égalisation des critères de la teinte qui modifie l'apparence visuelle des objets sur lesquels elle s'applique, une modification de la morphologie de la couleur. Son utilisation est très intéressante lorsque la figure est rendue complexe, soit parce que les objets présentent des valeurs combinées, c'est-à-dire des nuances de teintes qui correspondent à une double application, voire plus, de la variable visuelle *Valeur* (sur les teintes correspondantes)¹⁶⁹, soit parce qu'ils sont superposés, ou bien qu'ils se chevauchent comme dans le cas des flux.

La méthode est souvent mise en œuvre sur les cartes choroplèthes. Elle est aujourd'hui connue sous l'expression anglaise de *Value By Alpha (VBA) map*, introduite par A. Mac Eachren (1995) pour désigner la mise en œuvre de la variable visuelle *Transparence*. Il s'agit d'une méthode informatisée de *design* de la teinte formalisée en 2010 par R. E. Roth et al., en analogie avec l'utilisation de la *Taille* lorsqu'elle est mise en œuvre dans le cadre d'une anamorphose, ce qui explique la proximité des dénominations. La VBA est un clin d'œil à l'expression *Value By Area (VBA) map*, spécifique aux procédés d'anamorphoses scalaires qui déforment la morphologie des unités spatiales proportionnellement à la valeur d'un indicateur statistique. Ici, c'est plutôt la morphologie de l'arrangement des canaux (R, V, B) définissant la teinte perçue qui est déformée par rapport à la valeur d'un indicateur (une variable statistique ou spatiale).

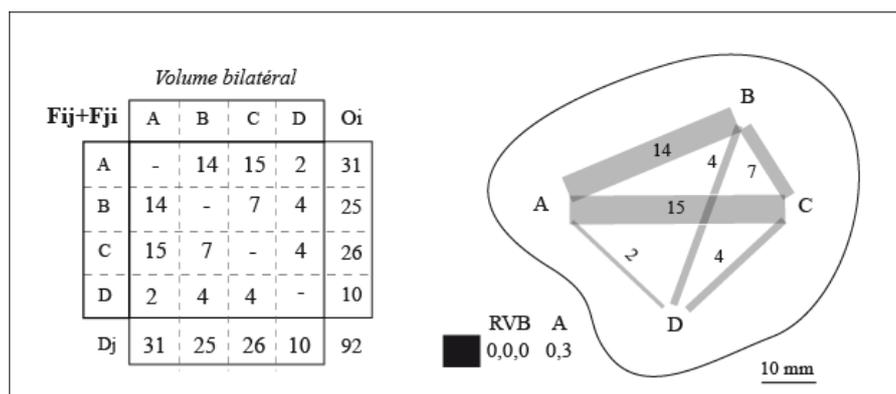
La mise en œuvre de la *Transparence* conduit donc à un ré-échantillonnage des niveaux – ou des classes – de teinte qui produit, au final, un double échantillonnage des mêmes données : au niveau de la valeur (*Valeur*) et au niveau de la teinte (*Couleur*), pour un effet visuel saisissant (voir Figure 5-6).

L'intérêt de l'utilisation de cette nouvelle variable visuelle est son applicabilité de deux manières : à chacune des unités spatiales qui caractérise la zone d'étude, de la même façon que la *Taille* pour l'anamorphose, ou bien aux *k* classes de valeurs, toujours comme la *Taille*. Une grande valeur alpha signifie que la teinte de l'objet est opaque, ou faiblement transparente et inversement. D'après les auteurs, la transparence de la teinte résout le dilemme de l'anamorphose (qui déforme la topologie et la morphologie des unités spatiales tout en autorisant une perception plus juste des motifs spatiaux), en égalisant parfaitement les teintes en fonction de l'objectif de la carte. Il est notamment possible d'agir sur la figure, en illuminant les unités spatiales qui présentent les valeurs les plus significatives (ou inversement), ce que les auteurs qualifient « *d'effet de projecteur* », ou d'ajuster la *Transparence* de manière à mettre en évidence un *effet de profondeur*. C'est cette dernière possibilité qui nous intéresse dans le cadre des flux.

L'application de la *Transparence* conduit, en effet, à la déformation de la teinte, de sa perception de manière à faire apparaître comme des gradients qui correspondent aux besoins de représentation de données qui sont incertaines (Mac Eachren, 1995) ou qui sont superposées, ce qui est le cas des données de flux. Elle permet, en effet, de bien gérer la superposition et le chevauchement des figurés, comme présenté sur la Figure 5-6 où la teinte des figurés correspond à 30% d'opacité.

¹⁶⁹ Cette combinaison correspond, par exemple, à la double sémiologie en aplats de deux séries de valeurs quantitatives représentées simultanément sur un même objet.

Figure 5-6. Illustration de la mobilisation du canal alpha de la teinte



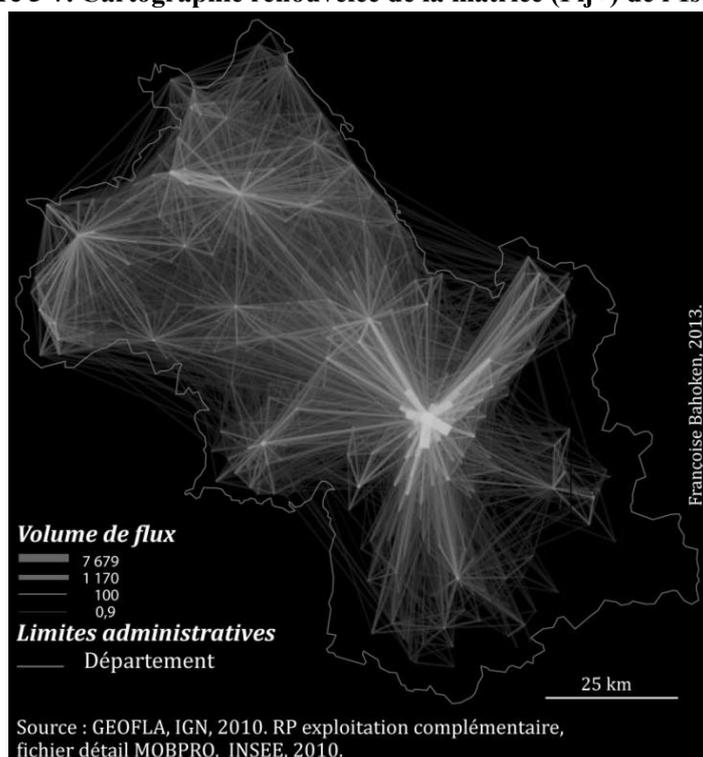
La largeur des bandes est proportionnelle à la valeur du flux

C'est pourquoi nous posons l'hypothèse que **cette utilisation de la *Transparence* de la teinte est susceptible d'améliorer le rendu des cartes de flux, tant sur le plan esthétique qu'au niveau de la perception des motifs.** Pour la valider, nous appliquons la *Transparence* à la carte de l'Isère (Fij⁺).

5.2.3 Application de la *Transparence* à la cartographie du volume bilatéral (Fij⁺)

L'utilisation de la *Transparence* est très récente. D'abord utilisée pour rehausser les cartes choroplèthes combinées, elle est ensuite mobilisée sur des cartes de visualisation des réseaux sociaux, caractérisées par un chevauchement important de figurés linéaires non valués. Ces cartes ne sont pas à proprement parler des cartes de flux qui suppose l'existence d'une quantité mesurable transférée. De plus, elles ne sont pas réalisées, dans l'ensemble, dans le respect des principes de sémiologie graphique y afférent.

La Figure 5-7 illustre la mise en œuvre de la *Transparence* à une cartographie de flux, illustrée par la matrice (Fij⁺) de l'Isère.

Figure 5-7. Cartographie renouvelée de la matrice (Fij⁺) de l'Isère

Le niveau de gris des figurés est obtenu automatiquement, par l'application d'un ré échantillonnage en fonction de la valeur du flux. Les figurés les plus courts apparaissent comme étant les plus opaques, en raison de la superposition de l'ensemble et non de leur teinte puisque le principal effet négatif de la *Transparence* réside dans la variation de la teinte lors des superpositions. Ainsi, plus le figuré est fin et plus sa teinte est transparente et sa disposition apparaît comme lointaine (les figurés sont tous placés sur le même plan), leur visibilité étant assurée par leur *longueur* graphique plus importante.

Cette carte illustre tout l'intérêt du recours à la *Transparence* de la teinte pour cartographier des matrices de flux symétriques, cette dernière améliorant d'ailleurs le rendu des cartes de flux de manière générale, tant au niveau esthétique qu'à celui de la perception des motifs. Elle diffère fondamentalement de la méthode classique réalisée sur la même carte, qui aurait eu pour résultat de générer une carte souffrant de *l'effet-spaghetti* et nécessitant l'application d'un critère de sélection des données.

Il est important de noter l'émergence d'un effet graphique particulier, *l'effet de profondeur*, rendu possible par l'utilisation de méthodes informatisées (voir Figure 5-7). Ce dernier est une forme *d'effet d'amalgame* correspondant au mélange d'informations de natures différentes (voir section 3.3.1, page 128). Sur le plan thématique, cet *effet de profondeur* suggère l'existence d'une composante d'échelle dans la distribution des valeurs de flux. Autrement dit, un système de mobilité à deux échelles qui traduit différentes *longueur de flux*, c'est-à-dire différentes distances parcourue par le flux. Etant donné que la prise en compte de *l'effet d'amalgame* permet de résoudre simultanément les deux composantes de *l'effet spaghetti* (la *surcharge graphique* et la *couverture spatiale*), on se rend compte qu'en réalité, c'est la prise en compte de la profondeur – par l'introduction de la *Transparence* – qui apporte une réponse d'ordre sémiologique au problème de *couverture spatiale*, les solutions d'ordre théoriques et mathématiques seront respectivement présentées dans les chapitres 6 et 7.

Ainsi, pour renforcer la perception des différentes catégories visuelles de liens représentés sur la Figure 5-7, afin qu'ils se distinguent par leur taille (*largeur*) ou par leur *longueur*, nous avons suivi la suggestion de W. Tobler (1987) de disposer les figurés qui présentent les valeurs les plus importantes au premier plan. A noter que l'auteur a également proposé de détourner les liens, avec un liseré clair (Tobler, 1987 : 159) à la manière des symboles proportionnels d'une carte choroplèthe, mais cette proposition reprise par D. Dorling (1991) ne nous est pas apparue efficace sur le plan esthétique.

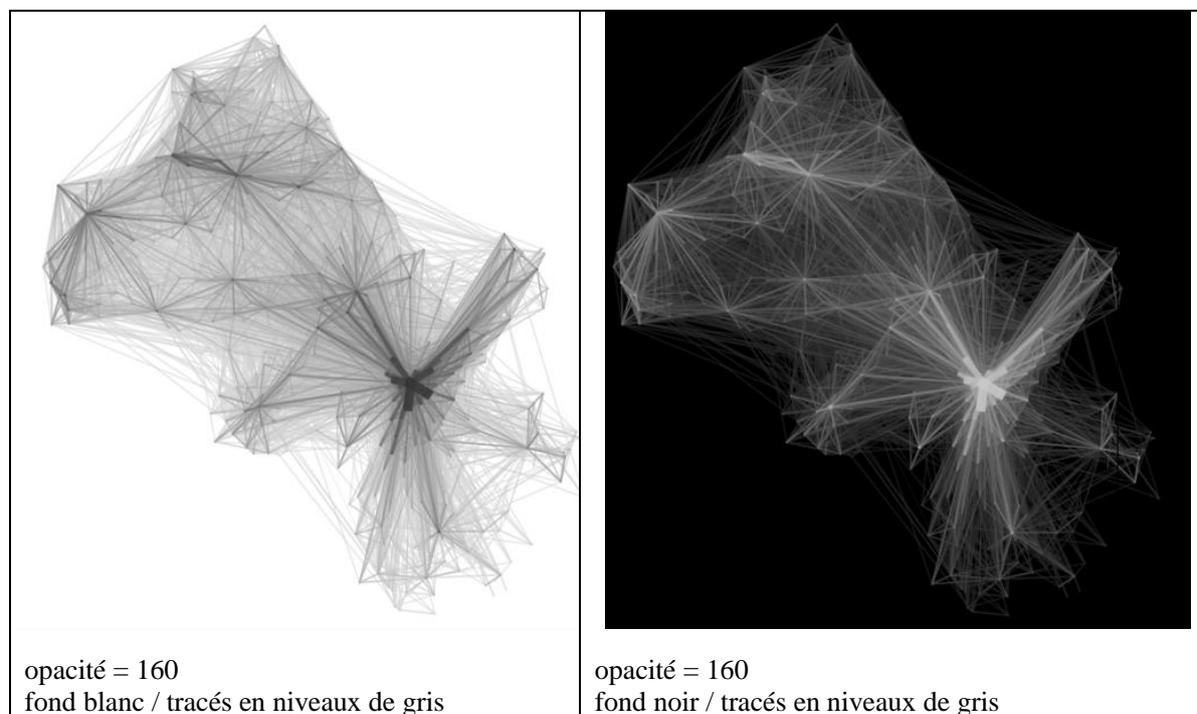
L'intérêt de l'utilisation de la *Transparence*, dans le contexte spécifique de la cartographie de flux, est donc avéré. **Il réside dans l'amélioration de la « saillance visuelle »¹⁷⁰ des figurés**, au sens de S.I. Fabrikant (2005), **aussi dans la perception de la masse de petits flux diffus qui présentent de faibles valeurs et qui ne seraient pas perceptibles autrement. La *Transparence* autorise, par conséquent, la représentation de toute l'information disponible dans la matrice, sans agir sur les valeurs mais sur la seule teinte.** La gestion simultanée de la disposition des liens, combinée à la *Transparence*, augmente à elle seule considérablement la séparation visuelle des figurés qui illustrent les flux et cela, d'autant plus que la figure est réalisée sur un fond noir dont nous avons vu qu'il augmente la perception des contrastes. Pour démontrer cette affirmation, nous présentons dans la sous-section suivante, une expérimentation qui agit sur la seule esthétique de la carte de la Figure 5-7.

¹⁷⁰ S. I. Fabricant introduit en cartographie la notion de « *saillance visuelle* », pour signifier l'adéquation entre le le phénomène représenté, les variables visuelles mobilisées et sa perception par un observateur. L'action sur la *Transparence*, telle que nous la mettons en œuvre sur des flux, permet ainsi de percevoir une part plus importante de figurés et d'information, tout en résolvant certains problèmes spécifiques de lisibilité.

5.2.4 Expérimentation sur l'intérêt esthétique et l'efficacité du fond noir et de la Transparence

L'expérimentation que nous réalisons consiste à modifier l'apparence de la carte du volume de flux (Fig¹) présentée sur la Figure 5-7, sur le seul plan de l'esthétique. Le résultat est présenté sur la Figure 5-8 : en noir sur fond blanc (à gauche) puis en blanc sur fond noir, pour une valeur d'opacité identique dans les deux cas.

Figure 5-8. Expérimentations sur l'esthétique du noir, à partir de la carte du volume de flux



Source : RP, exploitation complémentaire, fichier de détail MOBPRO, INSEE, 2010.

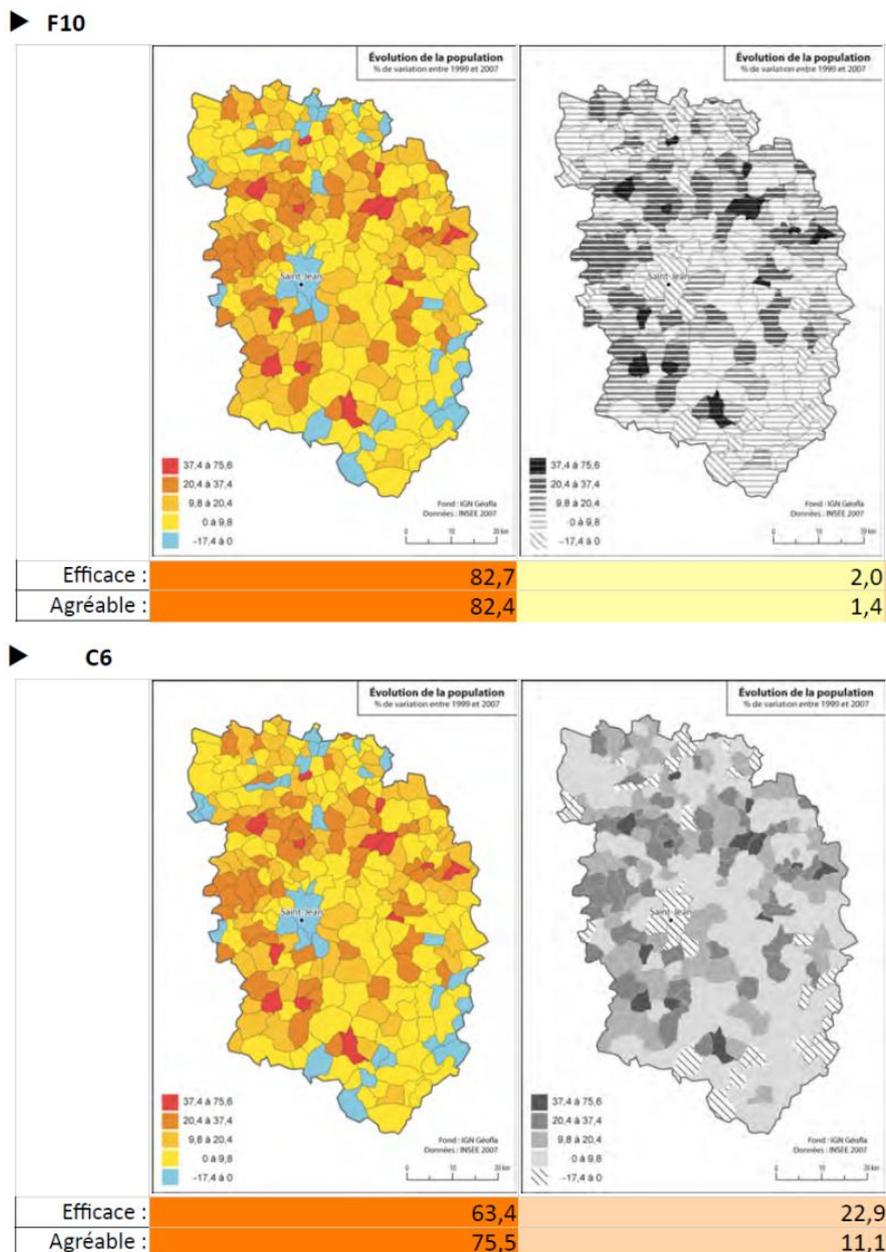
Nous rappelons avoir volontairement supprimé les symboles proportionnels illustrant ce volume aux lieux, car ils masquent l'information portée sur les liens, en raison de la manifestation de *l'effet de convergence* des plus gros flux vers des zones qui présentent, dans l'ensemble, les plus petites tailles. Réintégrer les symboles donnerait aussi une information erronée de la réalité car cela conduirait à réduire la *longueur* des bandes (en raison du jeu conjoint des *effets de convergence* et de *défonce graphique*). En effet, la gestion pratique de la convergence des flux entraîne une défonce des liens, c'est-à-dire une réduction de leurs dimensions pouvant être préjudiciable à la lecture de l'information représentée.

Sur l'expérimentation proposée dans la Figure 5-8, à partir de la Figure 5-7, nous ôtons l'ensemble des éléments d'habillage : le fond de carte des communes ainsi que les éléments de légende. La carte du volume de flux ainsi épurée est présentée sur la partie droite de la Figure 5-8 ; la partie gauche de cette même figure correspond à l'inversion de l'esthétique, dans le style classique du *noir-et-blanc*, pour les besoins de la comparaison. Dans les deux cas, les liens les plus courts sont placés au premier plan de la figure et leur degré d'opacité est plus élevé que les liens les plus longs.

De notre point de vue, ces deux cartes présentées sur la Figure 5-8, qu'elles soient réalisées en noir-et-blanc ou en blanc-et-noir, sont « agréables à regarder ». Cette affirmation, certes subjective, apparaît pourtant en faux avec l'une des conclusions de l'enquête réalisée par L. Jégou (2013b). En effet, « *Le noir et blanc est fortement rejeté [par les enquêtés], en efficacité comme en aspect esthétique* » (Jégou, 2013b : 44). « *Le rejet du noir et blanc est bien plus franc chez les anglophones. La méthodologie et*

les usages des francophones sont-ils encore trop marqués par le noir et blanc ? » (Jégou, 2013b : 47). Sans doute faut-il préciser que certaines questions de l’auteur consistaient à comparer l’esthétique de deux cartes choroplètes, selon qu’elles sont réalisées en couleur ou en noir et blanc. Les réponses aux questions F10 et C6 de l’enquête précitée ont particulièrement attiré notre attention. Dans les deux cas, la carte placée à gauche est réalisée en couleur, dans des tons de rouge-orangé vers le bleu ciel, pour les valeurs négatives (voir Figure 5-9).

Figure 5-9 .Les réponses aux questions F10 et C6 de l’enquête Cartographie et Esthétique (Jégou, 2013)

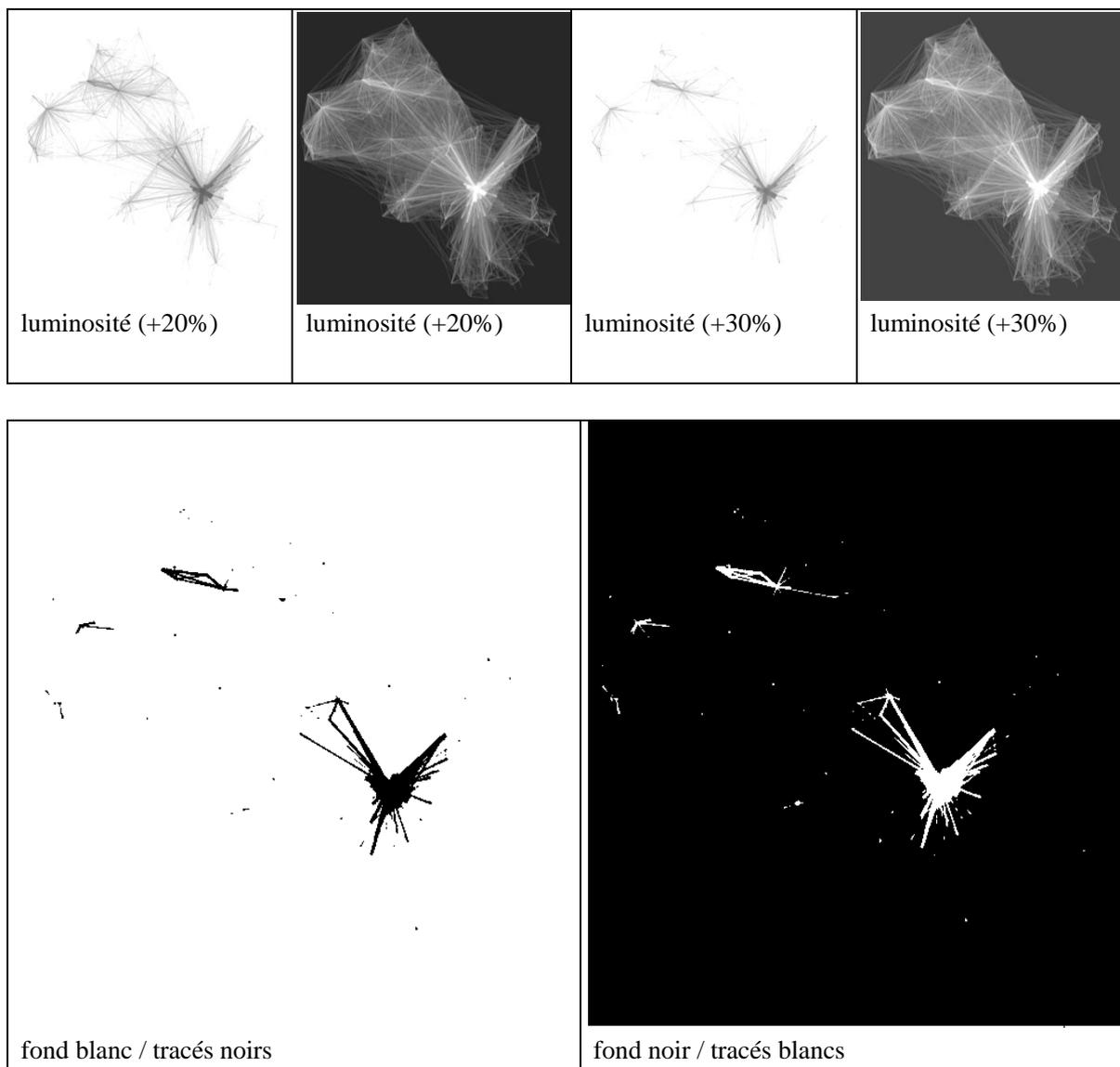


Cette disposition des cartes dans l’enquête est un choix parfaitement assumé par l’auteur. Dans les deux questions qui nous intéressent, la carte en couleur est perçue comme efficace et agréable, pour respectivement 82,7 % et 82,4 des réponses à la question F10 ; pour 63,4 et 75,5% des réponses à la question C6. Les deux cartes en *noir et blanc*, qui sont placées en vis-à-vis de la carte colorée, sont

composées, pour la question F10, d'hachures d'épaisseurs variables dont l'orientation varie en fonction de la polarité de la valeur ; pour la question C6, d'un mélange d'aplats en niveaux de gris pour les valeurs positives et de hachures pour les valeurs négatives. La carte réalisée avec des seules hachures en noir et blanc n'est absolument pas perçue comme efficace : seuls 2,0% des enquêtés la considèrent comme efficace et 1,4% comme agréable. Les résultats sont un peu plus nuancés pour la carte en noir et blanc de la question C6 : respectivement 22,9% et 11,1% des enquêtés la considèrent comme respectivement agréable et efficace. L'objectif de l'auteur étant aussi de comparer la perception des couleurs en aplat et des trames, les réponses seraient-elles les mêmes si la comparaison portait sur la couleur et le noir et blanc, sans hachures ?

Etant donné que l'image du volume des flux quotidiens de l'Isère doit, avant tout, être analysée au regard de ses principaux motifs spatiaux, nous avons réalisé le squelette des flux présentés sur la Figure 5-8. L'extraction de ce squelette, obtenu à l'aide d'un traitement d'image simple, est présentée sur la Figure 5-10.

Figure 5-10. Les principaux motifs du volume de flux de l'Isère en 2010



Source : RP, exploitation complémentaire, fichier de détail MOBPRO, INSEE, 2010.

Sur la Figure 5-10, les quatre images de la partie haute illustrent la manière dont le degré¹⁷¹ de luminosité et d'opacité influent sur la perception des flux, mais différemment en fonction de la teinte du fond. Lorsque celle-ci est claire, l'action sur la luminosité de la figure a un effet sensible : plus la luminosité est importante et plus l'information semble avoir été filtrée ou sélectionnée visuellement, les liens les moins gros disparaissant les premiers. La discrimination de l'ensemble des flux, pour une meilleure perception, passe ainsi par une action sur la luminosité des liens, mais aussi sur leur degré d'opacité qui se traduit par un aspect plus ou moins translucide. A l'inverse, lorsque le fond est noir, nous observons que l'action sur la luminosité n'affecte que peu, voire pas, la perception des figurés, ces derniers nous apparaissent d'ailleurs plus saillants, en contraste avec la teinte du fond. C'est en revanche, les différences d'opacité qui permettent de représenter l'ensemble des flux avec une bonne discrimination visuelle.

Les deux cartes de la partie basse de la Figure 5-10 ont pour objectif de démontrer que le changement de la teinte du fond n'a néanmoins, pas pour effet de modifier la perception des motifs principaux de la carte de flux et cela, afin que puisse être démontré l'intérêt de l'utilisation d'un fond noir.

Pour cela, il convient de comparer les cartes de la Figure 5-8 à celles qui sont situées dans la partie basse de la Figure 5-10. Celles-ci sont en tous points comparables puisque, d'une part, ces cartes correspondent à la même série statistique, elles sont donc strictement identiques du point de vue des données ; leur similitude porte d'autre part, sur les critères esthétiques qui ont conduit à leur élaboration : ces figures sont représentées avec un niveau de luminosité semblable. La comparaison visuelle des cartes met en évidence les flux qui se produisent entre les principales villes du département de l'Isère. On distingue au centre, le Y grenoblois, caractéristique d'une morphologie en cuvette coincée entre trois massifs alpins, qui s'étend de Vizille au sud jusqu'à la région de Moirans vers le Nord-ouest, celle de Crolles vers le Nord-est. Plus au Nord, les motifs concernent les flux qui se produisent dans le secteur de la ville nouvelle de l'Isle-d'Abeau incluant notamment Saint-Quentin-Fallavier ; la région de Vienne se distingue aussi au nord-ouest du département et cela, que le fond de la carte soit teinté de Blanc ou de Noir. Si nous n'observons pas de modifications dans la perception des principaux motifs de ces flux, lorsque nous comparons visuellement les cartes de la Figure 5-8 avec celles de la Figure 5-10, nous percevons tout de même davantage de détails lorsque le fond est Noir et les figurés blancs : ceci tient probablement à un degré de séparation plus important des figurés lorsque le fond est Noir. Parce que ce constat est vraisemblablement fortement subjectif, nous attirons l'attention sur le fait que ces cartes ayant été réalisées sur un écran, des différences de perception des nuances – et plus tard dans le texte de la brillance des couleurs – sont possibles, dans le cas où l'impression présent du document n'aurait pas été réalisée en couleur, à un niveau de résolution suffisant (>300dpi).

Cette première analyse visuelle comparative des principaux motifs des flux, selon que le fond est *Blanc* ou *Noir*, nous conduit à deux conclusions : la modification de la teinte du fond n'engendre pas de différences dans la perception des contrastes des motifs du volume de flux représenté, elle n'altère pas leur compréhension ; l'harmonie colorée étant par ailleurs préservée. L'action réalisée sur l'opacité des liens (Figure 5-10, en haut) permet de révéler la masse de petits flux diffus qui étaient jusque là *invisibles*. Elle procure aussi un effet visuel d'emboîtement, lié au placement des plus petits liens en arrière-plan, qui met en évidence une hiérarchie des flux. Celle-ci suggère une mobilité qui se produit à différentes échelles spatiales : localement, c'est-à-dire à courte distance mais aussi à moyenne distance (nous réservons l'expression de longue distance aux flux d'envergure nationale). Les lieux les plus attractifs apparaissent également plus saillants dans ce système de mobilités et c'est comme s'ils étaient placés au premier plan, ce qui n'est pas le cas.

¹⁷¹ Les critères utilisés pour modifier la luminosité et l'opacité des figures sont purement subjectifs. Ils ont été définis visuellement et permettent d'aboutir à un degré d'opacité de 100.

L'équivalence cognitive entre le *Noir* et le *Blanc* du fond démontre que le choix de l'une ou de l'autre teinte est indépendant des considérations théoriques. Il n'est lié qu'aux aspects esthétiques de la figure, ce qui a pour conséquence que nous pouvons tout à fait substituer l'un à l'autre et inversement, pour les motifs des flux. Nous faisons donc le choix de la version de la carte de flux en *blanc-et-noir*, pour poursuivre l'examen de la *Couleur*.

La méthode autorise donc, de notre point de vue, la représentation d'une quantité d'information plus importante et la perception d'une information qui était auparavant invisible. Sur le plan graphique, la *Transparence* permet de ne pas représenter l'effet des flux du point de vue des lieux, par des points proportionnels. On voit bien, de notre point de vue, sur la Figure 5-7, quels sont les lieux dont le volume de flux est le plus important.

La variable visuelle *Transparence* s'applique bien au cas des matrices de flux denses et symétriques, (réalisées avec des bandes) où *l'effet de convergence* des flux peut ne pas être représenté, tout en étant visible (sa suppression renforçant en fait la perception des motifs). La *Transparence* n'est toutefois pas suffisante pour représenter les matrices antisymétriques (F_{ij}), car elles nécessitent l'utilisation d'un figuré orienté, une flèche par exemple.

Dans la section suivante, nous revenons sur le cas classique de cartographie du (F_{ij}), afin d'examiner la manière dont une nouvelle utilisation de la *Couleur*, en tant que variable visuelle, est susceptible de renouveler la sémiologie cartographique des matrices de flux antisymétriques.

5.3 Le renouvellement de la variable visuelle *Couleur*, l'exemple du (F_{ij})

Dans la section précédente, nous avons mobilisé la couleur en tant que teinte, pas en tant que variable visuelle. Etant donné que nous avons montré que la cartographie du *volume bilatéral de flux* (F_{ij}^+) ne nécessitait pas l'usage de la couleur, c'est donc à partir de la matrice du bilan de flux (F_{ij}) que nous allons examiner la mise en œuvre de la *Couleur*.

La *Couleur* est mise en œuvre de deux manières en cartographie statistique, selon qu'elle est utilisée seule, comme un petit nombre de teintes représentant des catégories différentes entre elles, ou en association avec la variable visuelle *Valeur*, sous la forme de nuances de teintes. Dans le premier cas, la *Couleur* est appliquée à des valeurs nominales, caractéristiques qualitatives des objets. C'est le cas, par exemple, de la cartographie de modes d'occupation des sols variés, en fonction d'une typologie de type 'espaces en eau' ; 'espaces urbanisés', 'espaces agricoles' qui prennent des teintes distinctes. Le second cas d'utilisation classique de la *Couleur* est mis en œuvre sur des variables quantitatives continues qui ont été discrétisées en k plages distinctes de valeurs, sur lesquelles s'applique la *Valeur*. Il s'agit d'une variable visuelle de l'image, fortement sélective, qui entraîne l'utilisation d'un camaïeu de k tons (un par plage de valeurs) ordonnés d'une même teinte.

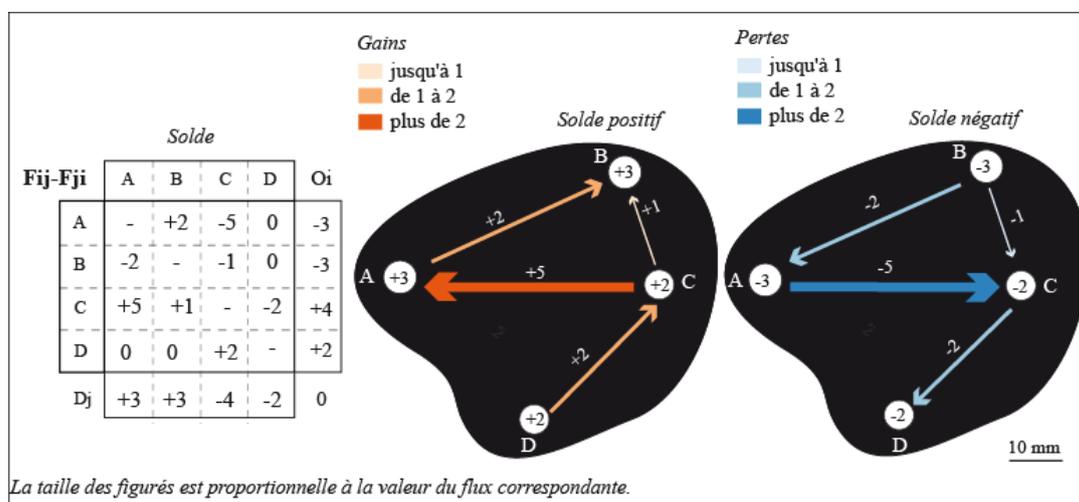
C'est parce que cette teinte est souvent colorée que les variables visuelles *Valeur* et *Couleur* sont généralement associées, pour combiner leurs propriétés respectives d'associativité et de séparation. Cette combinaison se justifie d'autant plus que la série analysée présente des valeurs positives ou négatives, dont la concomitance justifie l'utilisation du contraste de couleurs, de la sémantique de l'opposition colorée aussi qualifiée d'inversion ou de divergence des valeurs.

5.3.1 La sémantique de l'opposition colorée

Le contraste coloré est spécifique à l'utilisation de la *Couleur*. Il permet d'associer des teintes *chaudes-froides* ou encore *claires-obscur* pour signifier une opposition. En effet, pour représenter le changement de polarité des valeurs, il convient d'utiliser des teintes opposées pour distinguer les valeurs en fonction de leur signe. L'utilisation de cette sémantique d'opposition permettant de renforcer la propriété de séparation de la *Couleur*.

En cartographie statistique, cette utilisation combinée des variables visuelles *Valeur* et *Couleurs opposées* est toujours et exclusivement mise en œuvre lorsque l'implantation des objets est surfacique et ponctuelle, plus rarement linéaire. Lorsqu'elle est mise en œuvre sur des figurés en implantation linéaire, le principe d'application de la *Couleur* est toujours classique, il est illustré sur la Figure 5-11.

Figure 5-11. Illustration de la mise en œuvre, sur des flux, des nuances de teintes opposées



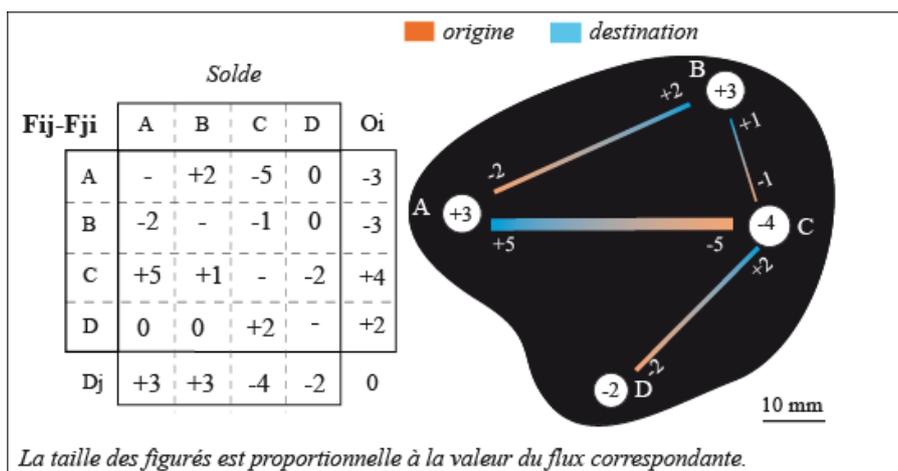
Ce procédé classique illustré sur la Figure 5-11, consiste à affecter les bandes d'une coloration discrétisée à la manière des points, pour renforcer l'appartenance à des classes de valeurs issues d'une discrétisation. Il n'est pas dénué d'intérêt, lorsque la matrice est dense puisqu'il permet de ne réaliser que k classes de tailles des figurés, une pour chaque nuance de teinte, et non une représentation proportionnelle de l'ensemble des figurés. Un ajustement efficace de la taille associée à chacune des classes présente aussi l'avantage de résoudre les problèmes liés à un étalement important des valeurs statistiques.

L'association *Valeur* et *Couleurs opposées*, telle qu'elle est mise en œuvre en cartographie statistique, ne semble pas faire l'objet, à notre connaissance, d'une application pour les flux, c'est-à-dire au niveau des liens représentés par des bandes de *Taille* variable. Lorsque la matrice est dense et orientée, le transfert de la sémantique d'opposition au niveau des liens permet de conserver la hiérarchie des lieux et de résoudre le problème de *l'effet de convergence* des flux, sans souffrir de la résolution des contraintes graphiques liées à l'utilisation de la flèche.

Dans un style en blanc-et-noir, les figurés étant portés en blanc, il est dès lors possible de supprimer les têtes de flèches afin de ne pas alourdir la figure, puis de transférer l'opposition colorée au niveau des bandes, comme illustré par la Figure 5-12, située page suivante.

On y observe que **l'utilisation d'un gradient de couleurs opposées pour représenter le solde bilatéral revient à fusionner en un seul objet, les deux précédentes images des flux correspondant aux cartes des créances et des dettes, respectivement aux valeurs de flux positives et négatives.** La matrice étant antisymétrique, nous avons vu (section 3.3.3) qu'il était nécessaire de sélectionner les valeurs (positives ou négatives) à représenter, soit représenter une partie de l'information.

Figure 5-12. Le renouvellement de la sémiologie classique du solde bilatéral (Fij-)



L'intérêt de l'utilisation de l'opposition colorée concerne, en effet, spécifiquement la cartographie des matrices antisymétriques dont le résultat produit désormais une image unique. Elle présente une seule relation dont la particularité est d'être *antisymétrique* : il s'agit bien d'une même quantité échangée entre deux lieux, mais sa signification varie en fonction du point de vue (ou du lieu) duquel on l'observe. En filant la métaphore financière, c'est comme si le lieu fonctionnait comme une banque qui prête à certaines banques et emprunte aux autres, la convergence des teintes rouges et bleues permettant de qualifier les perdants et les gagnants d'où le fait que l'on propose de représenter les créances et les dettes sur la même carte.

L'utilisation d'une opposition colorée au niveau des bandes qui décrivent les flux est fructueuse. Elle présente plusieurs avantages qui sont, en premier lieu, d'évincer la flèche typique pour les cartes de flux orientés, tout en conservant le sens des flux par le système d'opposition colorée. Sur la Figure 5-12, les teintes orangées illustrent l'origine des flux : l'émissivité des lieux, tandis que celles qui sont bleutées indiquent la destination ou la réceptivité des flux (les créances ou bien les dettes). Cette méthode conduit aussi à résoudre le problème de la représentation de *l'effet de convergence* puisque nous avons observé, sur des données réelles, qu'il n'était plus nécessaire de représenter les soldes des lieux, aux côtés des liens, tout en conservant l'idée de lieux d'accumulation et de solde. Elle présente toutefois une limite non négligeable qui est la suivante : la lisibilité des couleurs n'est assurée que lorsque le trait est suffisamment large.

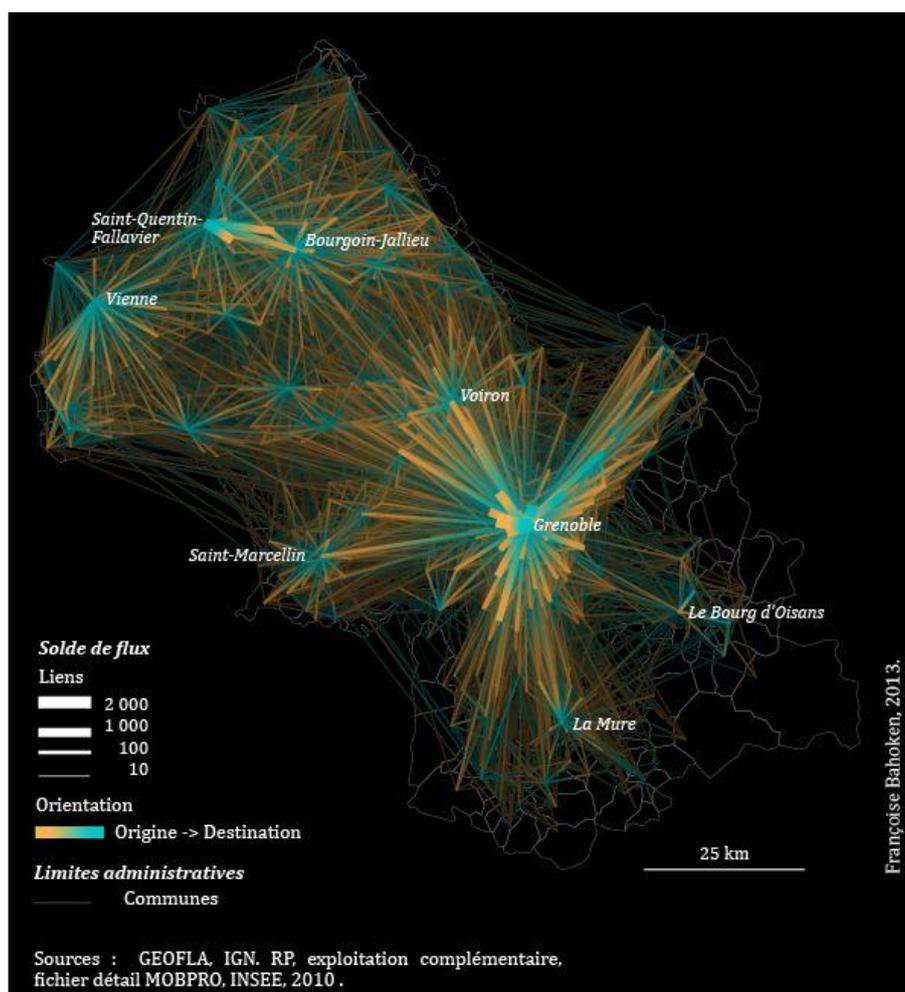
5.3.2 Application empirique de l'opposition colorée à la cartographie du solde bilatéral (Fij)

Nous illustrons la mise en œuvre de la sémantique de *l'opposition colorée* à la cartographie de la matrice (Fij) de l'Isère, qui correspond au cas réel¹⁷² d'application de la variable visuelle.

Le résultat est présenté sur la Figure 5-13, dans l'esthétique du *blanc-et-noir*. Cette figure mobilise ainsi la variable visuelle *Taille* de façon classique, mais la *Couleur* y est mise en œuvre de trois manières : en tant que teinte, au niveau esthétique du style de carte qui est réalisée sur un fond sombre ; au niveau sémiologique en mobilisant d'une part, la *Transparence* pour cartographier le solde bilatéral de la même manière que le volume bilatéral et d'autre part, un gradient de couleurs opposées sur les liens.

¹⁷² Contrairement à la pratique usuelle en Informatique graphique, l'attention est attirée sur le fait que l'application du principe de l'opposition colorée à (Fij) est erronée, si l'on ne tient pas compte de la thématique du flux, pour la raison suivante. (Fij) étant asymétrique, sa cartographie nécessite le recours à deux figurés, dont les largeurs diffèrent nécessairement (sinon la matrice serait symétrique). Ces deux figurés symbolisent (Fij) et (Fji), avec (Fij) ≠ (Fji), ce qui interdit leur réduction à un figuré unique. Ce constat renforce par ailleurs la nécessité d'une réduction de la complexité de la cartographie directe de (Fij), par une décomposition de l'information.

Figure 5-13. Cartographie renouvelée de la matrice (Fij) de l'Isère



Il est important de noter sur cette Figure 5-13 qu'il n'est plus nécessaire de représenter les flux du point de vue des lieux (par exemple, par des symboles proportionnels au solde des flux aux lieux), comme sur la Figure 3-19 (page 133), pour une lecture liens-lieux de ces flux. On observe, en effet, que *l'effet de convergence* des flux pour illustrer l'attractivité (ou l'émissivité) des lieux est assurée par la seule représentation des liens.

Cette éviction de la représentation ponctuelle des lieux se justifie par un aspect d'ordre sémiologique, aussi par la thématique de la carte. Les données décrivant des mobilités quotidiennes, il n'y a pas lieu d'analyser la carte du bilan net du point de vue des gains ou des pertes. Ce qu'il faut y voir, c'est plutôt la mise en évidence de cycles de présence diurne et nocturne en fonction de l'activité. En effet, les lieux de convergence des bandes bleues correspondent aux pôles d'emplois de l'Isère où la population est augmentée pendant les heures de travail, tandis que ceux des bandes orange correspondent aux lieux résidentiels où la population est augmentée la nuit et le week-end. Les liens courts mettent en évidence le cœur du Y grenoblois, coincé entre les massifs du Vercors à l'ouest, de la Chartreuse au nord, de Belledonne et de l'Oisans à l'est. Il en est de même pour les principaux axes de circulation qui participent, très probablement, de ces mobilités quotidiennes : la rocade sud A480, qui dessine une première couronne fonctionnelle du sud-ouest au nord-est de la ville de Grenoble, aussi l'autoroute A48 vers Voiron, puis vers les villes de Bourgoin-Jallieu et de l'Isle-d'Abeau situées plus au nord du département ; l'autoroute alpine A41 vers la Savoie, en passant par Pontcharra et

Allevard. Les pôles secondaires attractifs de ce système de mobilités quotidiennes sont ainsi également perceptibles sur les deux cartes, alors qu'il ne s'agit pas de la même représentation.

La *Transparence* et l'*opposition colorée* se prêtent bien aux cartes qui décrivent de manière isolée le *solde bilatéral* (F_{ij}) des flux, soit la partie dynamique de l'information disponible dans (F_{ij}). Néanmoins, on peut considérer que la représentation de ces matrices antisymétriques souffre de deux maux : d'une part, le fait que ne soit pas également représentée, sur la même carte, la partie complémentaire de l'information (F_{ij}^+) et, d'autre part, qu'il faille tout de même produire deux cartes distinctes : l'une pour (F_{ij}^+) et l'autre pour (F_{ij}) pour cartographier *in fine* les mêmes données initiales (F_{ij}). Pour y remédier, il pourrait également être intéressant d'examiner le cas de l'asymétrie des flux.

Nous présentons dans la section suivante une validation empirique des deux solutions de cartographie proposées pour (F_{ij}^+) et (F_{ij}) sur une matrice de flux commerciaux qui s'exprime à l'échelle mondiale.

5.4 Validation empirique : l'exemple de flux commerciaux internationaux

Cette section a pour objectif de valider les solutions sémiologiques proposées qui conduisent au renouvellement de l'utilisation de la *Couleur* : par une action sur la *Transparence* (Mac Eachren, 1995), applicable à toutes les matrices, et par le transfert du principe de l'*opposition colorée* pour cartographier les matrices de flux antisymétriques (les transferts nets bilatéraux). Pour cela, nous avons recours à une matrice de flux internationaux mondiaux de marchandises (2004-2006) exprimées en milliards de dollars américains. Les données proviennent de la base de données CHELEM, préparée par C. Grasland, dans le cadre du groupe de travail consacré à l'analyse des flux et des réseaux (Work Package 5) du programme de recherche européen Eurobroadmap¹⁷³ (2009-2011). Nous présentons dans la première sous-section les données disponibles dans cette matrice, puis la mise en œuvre des solutions proposées à des matrices de type (L_{ij}), (F_{ij}^+) et (F_{ij}) de flux internationaux commerciaux.

5.4.1 Présentation des données

La base de données CHELEM est composée de trois bases que sont CHELEM-CIN, CHELEM-PIB et CHELEM-BAL, portant respectivement sur le commerce international, la population, le Produit intérieur brut (PIB) et les taux de change ; et sur la balance des paiements. La matrice dont il est question ici est issue de la base CHELEM-CIN.

Trois années (2004, 2005 et 2006) ont été agrégées sur une base temporelle, géographique et thématique, couvrant l'ensemble des produits échangés entre 91 unités spatiales (pays ou groupes de pays), en valeur (dollars américains courants, US\$).

Sur le plan temporel, les données ont été agrégées en une période tri-annuelle pour des raisons de robustesse statistique, en particulier éviter les fluctuations exceptionnelles. La matrice (2004-2006) que nous avons utilisée décrit ainsi la moyenne mobile des valeurs de flux internationales sur les trois années.

Sur le plan thématique, les données sont disponibles en 71 catégories de produits échangés, complétées par une catégorie « *non ventilé* », éventuellement regroupées en différentes nomenclatures¹⁷⁴. Pour notre part, si nous avons fait le choix dès le chapitre 1 de n'analyser qu'une seule matrice (F_{ij}) portant sur un ensemble, nous pourrions toutefois mobiliser en tant que de besoin

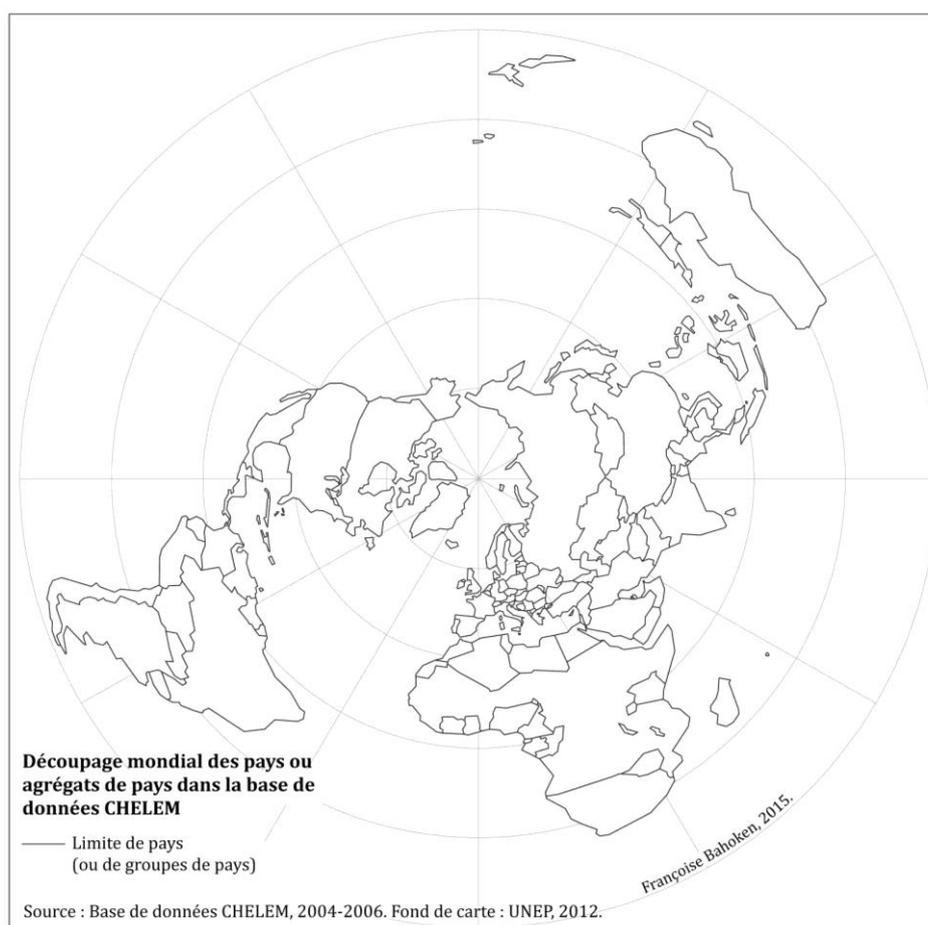
¹⁷³ Les travaux menés dans le cadre de ce projet, sur l'analyse des flux et des réseaux, sont disponibles à l'adresse suivante : http://www.eurobroadmap.eu/Flows_and_Networks [vérifiée le 02/08/2015]

¹⁷⁴ « *La nouvelle base propose au choix les 43 postes de la nomenclature GTAP échangés et les 147 postes CITI révisée 3 à 4 positions échangés, ainsi que des agrégations en 11 filières, 10 sections, 6 stades et 4 secteurs, et nouvellement 4 niveaux technologiques regroupant 20 agrégats* » (Cepii, 2008 : 10)

des sous-matrices ventilées en huit catégories, en sus de la variable « Total » (TOT) des marchandises puisque la matrice CHELEM est catégorielle. Les types de marchandises qui sont disponibles pour notre analyse sont l'ensemble des flux d'énergie (ENE), d'électronique (ELE), de minerais (MIN), de produits agricoles (AGR), de textile (TEX), de biens d'équipements (EQU), de chimie (CHE) et de produits divers (MSC).

Sur le plan géographique, le fond de carte mondial associé à ces flux internationaux de marchandises est formé de zones hétérogènes d'origines variées, comme illustré par la Figure 5-14.

Figure 5-14. Le découpage du monde, selon CHELEM



Le découpage mondial, selon CHELEM, est composé d'un mélange de pays et de groupes de pays. Dans le premier cas, il s'agit de l'un des 82 pays qui a fait l'objet de la collecte et dans le second cas, soit de l'une des 12 zones d'agrégats de pays, soit de la zone « divers » qui concerne « [...] *les flux sans affectation géographique connue* » (CHELEM, 2008). Les agrégats disponibles regroupent de petits pays (de l'Amérique latine, d'une partie de l'Afrique ou de l'Asie du Sud-est) ayant un volume commercial faible ou peu stratégique.

L'intérêt du recours à cette matrice de flux internationaux réside dans le fait qu'elle diffère fondamentalement du cas de l'Isère : les valeurs portent sur des flux internationaux et non sur des flux internes et sont pluriannuelles. Elles ont fait l'objet d'une agrégation temporelle et correspondent à une valeur monétaire moyenne des flux observés sur la période (2004-2006). La matrice décrit donc un échange commercial en valeur (US \$) pouvant être assimilé à un flux financier de marchandises.

Cette matrice est de type (Fij). Compte tenu de la thématique du flux – des échanges commerciaux –, elle nécessite d'être analysée sous une forme décomposée, comme présentée dans le Chapitre 3. (Fij)

étant, en effet, nécessairement différent de (F_{ji}) , car les flux d'exportations et d'importation qu'elles décrivent respectivement correspondent à une relation économique d'offre et de demande entre des pays ou des groupes de pays. La position de ces derniers, d'acheteur ou de vendeur, varie en fonction du sens de la relation économique. Ces flux peuvent dès lors être analysés du point de vue du total des flux (F_{ij}) , qui correspond aux exportations ; du point de vue des importations (F_{ji}) ; du volume bilatéral d'échanges (F_{ij}^+) ou du solde bilatéral (F_{ij}) . Ces différentes solutions sont proposées dans la sous-section suivante.

5.4.2 Les solutions proposées

L'application des propositions de ce chapitre à cette matrice des flux commerciaux est présentée sur les trois figures suivantes, respectivement consacrées à la cartographie de (L_{ij}) , de (F_{ij}^+) puis de (F_{ij}) selon deux esthétiques : d'abord en *noir-et-blanc* puis en *blanc-et-noir*. L'objectif de cette double présentation est de valider l'intérêt du recours au fond noir pour représenter une masse importante de figurés enchevêtrés et superposés, en mobilisant la *Transparence* et le principe de l'opposition colorée de la *Couleur*, au niveau des liens. Le fond noir permet, en effet, de mieux discriminer les petits flux diffus qui sont caractéristiques de ces matrices denses. Il renforce la perception des contrastes de ces cartes, pourtant complexes par la quantité de figurés qui y est représentée, grâce au recours à des teintes lumineuses.

L'objectif est aussi de comparer les résultats obtenus par l'application des solutions proposées, avec ceux de travaux réalisés sur les mêmes données, publiés respectivement par : Y. Richard et C. Zanin dans la revue *Géocarrefour* en 2009 ; C. Grasland et G. Van Hamme dans *l'Espace Géographique*, en 2010. Ces deux groupes d'auteurs soulignent la structure triadique de ces flux commerciaux et l'organisation spatiale de type centre-périphérie qui découle de ces échanges.

La Figure 5-15 présente une des cartes proposée par Y. Richard et C. Zanin, elle décrit des effets de barrières au commerce international.

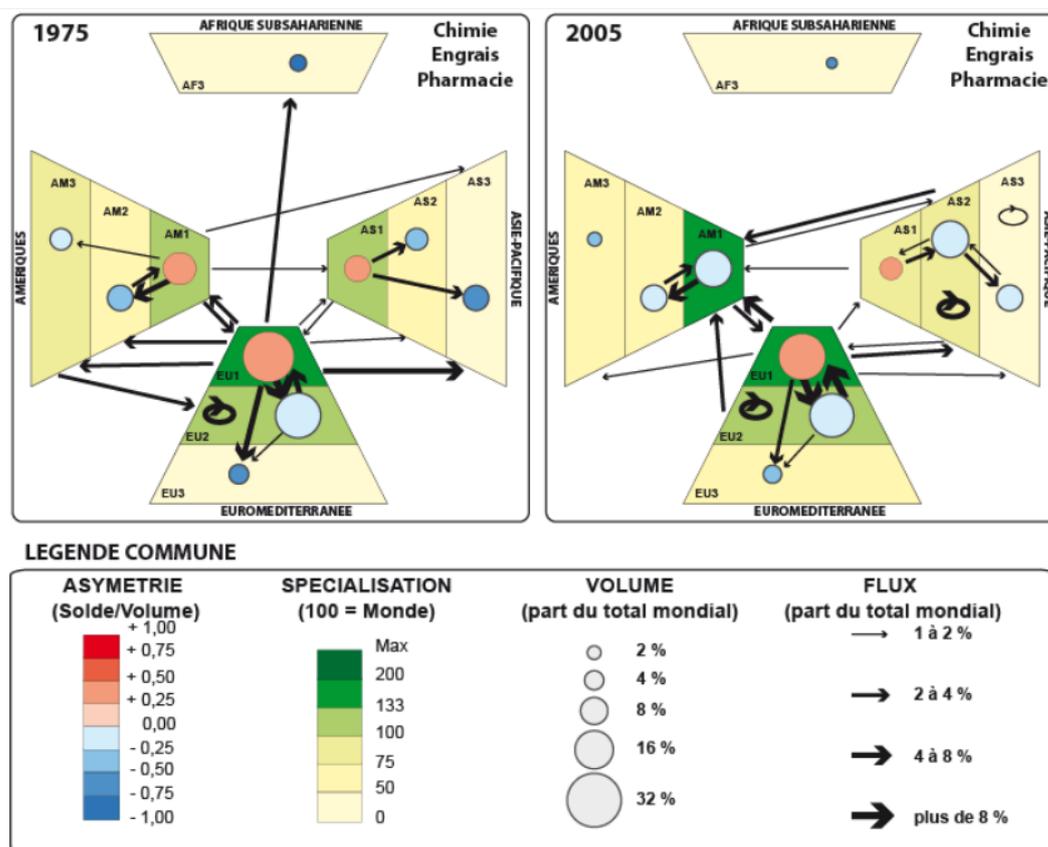
Figure 5-15. Cartographie de liens commerciaux internationaux (Zanin, Richard, 2009)



Sur la Figure 5-15 on devine l'organisation globalement triadique de ces flux commerciaux. D'après les auteurs, cette « carte montre plusieurs choses : (1) il n'y a pas de résidus négatifs entre des pays très proches appartenant à la même région ; (2) les effets de barrière (résidus négatifs) sont plus nombreux entre l'Europe et l'Asie de l'Est et du Sud-Est et entre l'Europe et l'Amérique du Nord, qu'entre l'Amérique du Nord et l'Asie de l'Est et du Sud-Est ; (3) il n'y a pas d'effets de barrière entre des pays du Sud ; (4) les effets de barrière entre des pays du Sud et des pays du Nord sont rares et, lorsqu'ils existent, ils concernent des pays qui n'appartiennent pas au même fuseau terrestre (on voit par exemple des résidus négatifs entre des pays en développement de l'Asie du Sud-Est et des pays de l'Europe ». Nous allons tenter de percevoir ces motifs dans la première des deux séries de cartes suivantes décrivant l'ensemble des flux mondiaux, dans l'esthétique du *blanc-et-noir* et du *noir-et-blanc*, sous de manière décomposée : les liaisons commerciales existantes (Lij) ; les volumes bilatéraux d'échanges commerciaux (Fij⁺) ; les soldes bilatéraux d'échanges commerciaux (Fij⁻) ; Les flux d'exportation (Fij) – ou des flux d'importation (Fji) – commerciaux mondiaux.

L'article de C. Grasland et G. Van Hamme vise notamment à mettre en évidence la pertinence de l'approche centre-périphérie pour analyser l'évolution géoéconomique mondiale (entre 1975 et 2005) et cette même structuration spatiale triadique. Les auteurs focalisent en particulier leur attention sur les trois macro-régions mondiales que sont les Amériques, l'Asie et Euro-méditerranée qu'ils ont définies par ailleurs. Ces régions, également structurées selon une logique centre / périphérie, font l'objet de représentations graphiques. La figure suivante en est l'une d'elles qui porte sur les produits chimiques, engrais et pharmacie (CHE).

Figure 5-16. Représentation graphique des échanges de produits chimiques entre grandes régions mondiales (Grasland, Van Hamme, 2010)



Auteurs : Grasland C., Van Hamme G., 2009

Source : CHELEM-CEPII

Figure 5-17. Cartographie des liaisons commerciales (Lij) existantes

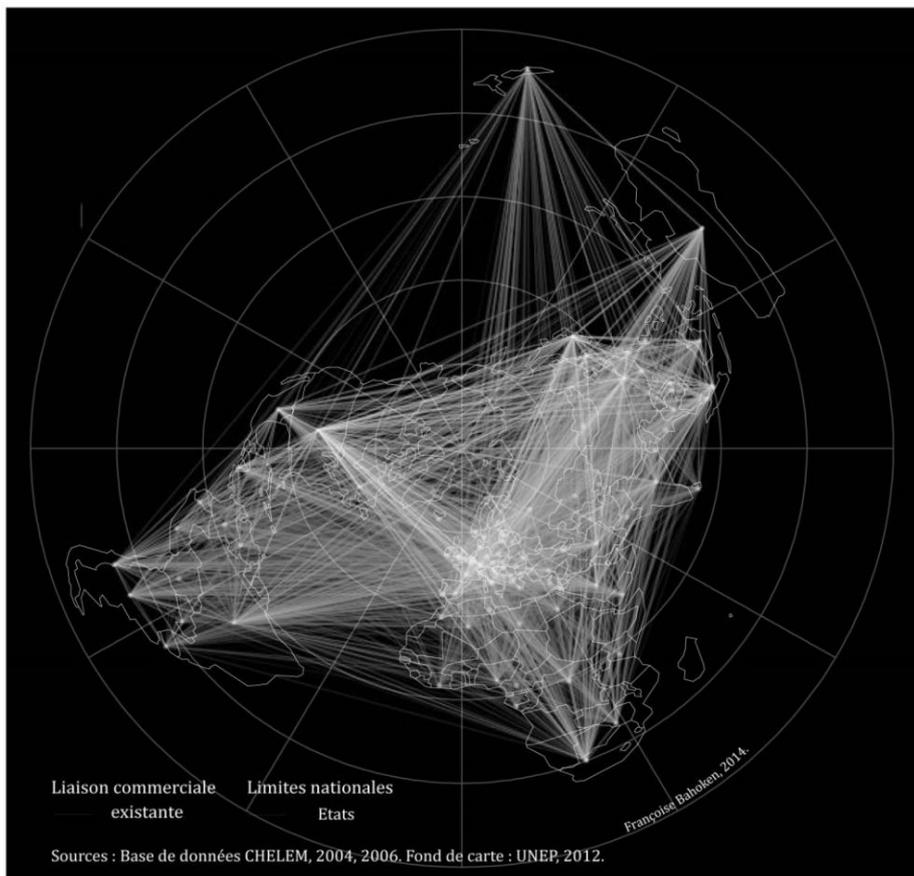
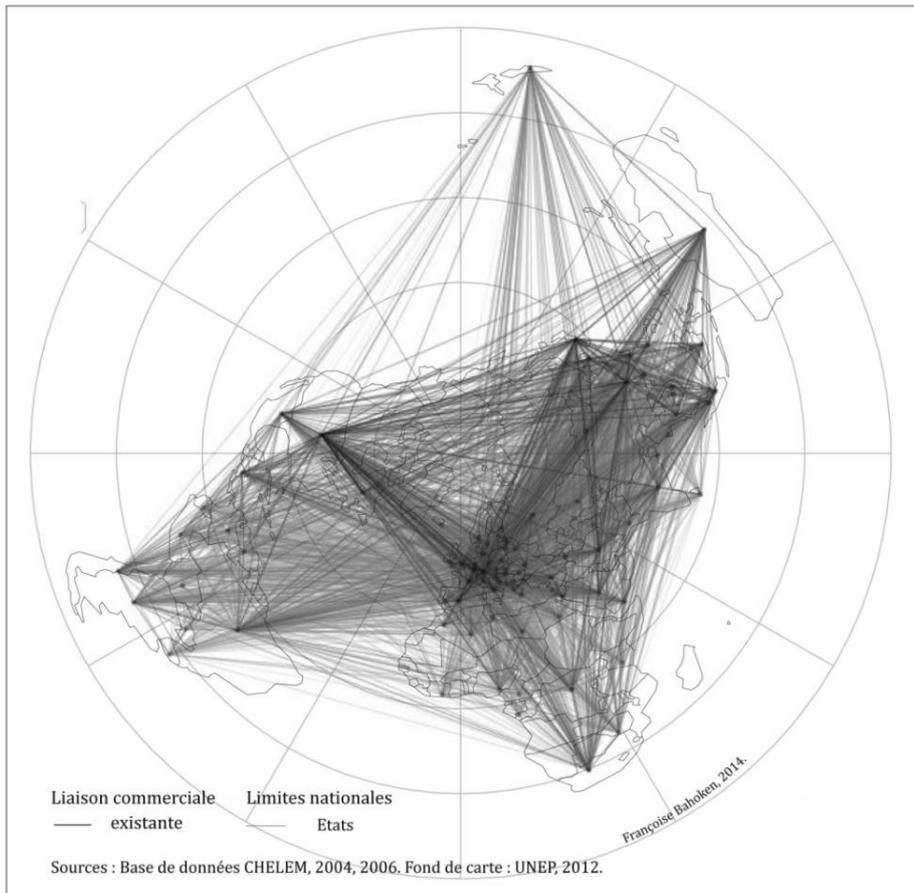


Figure 5-18. Cartographie des volumes d'échanges commerciaux (Fij⁺)

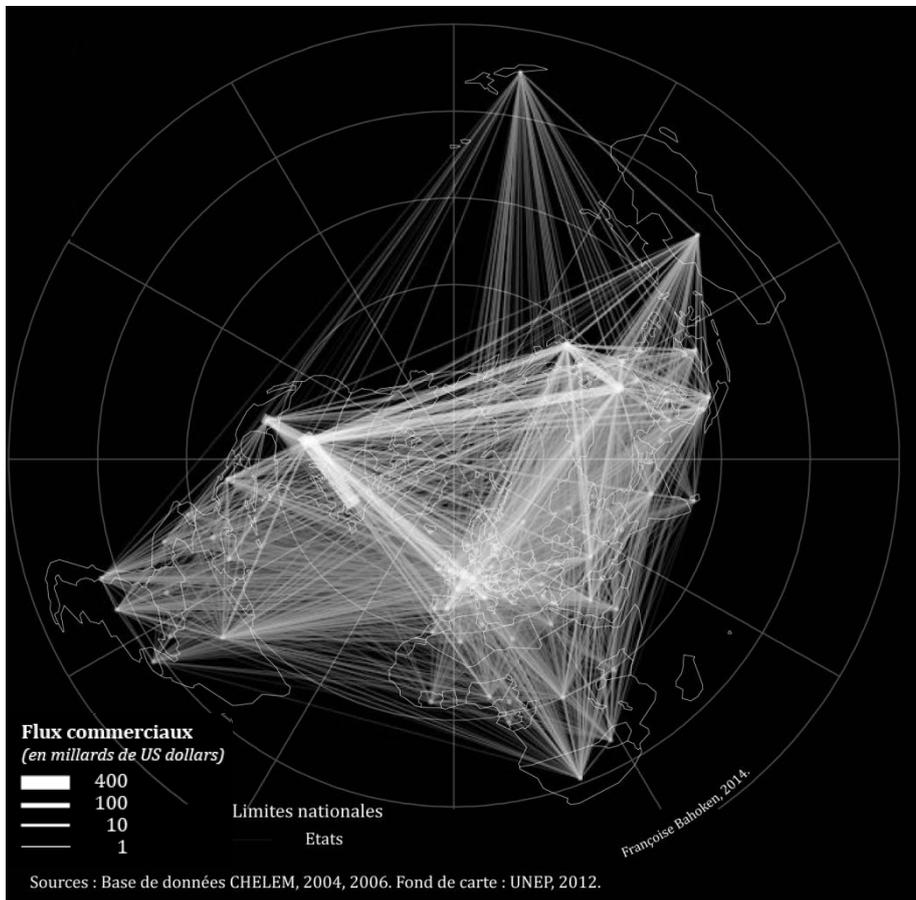
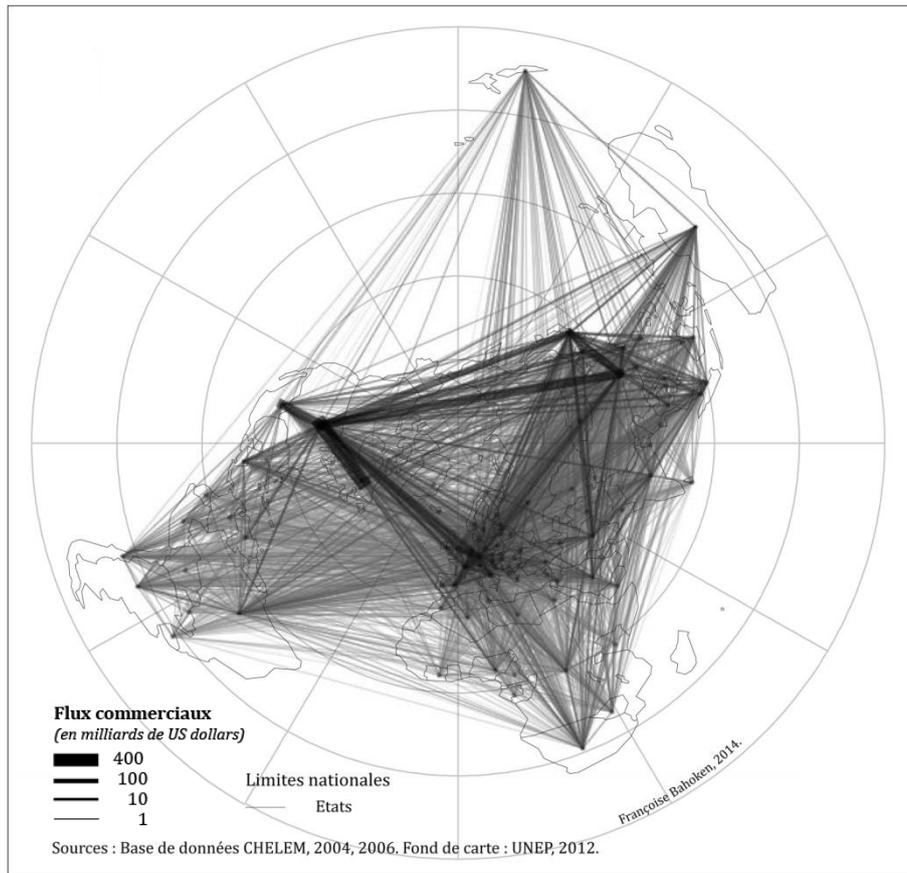
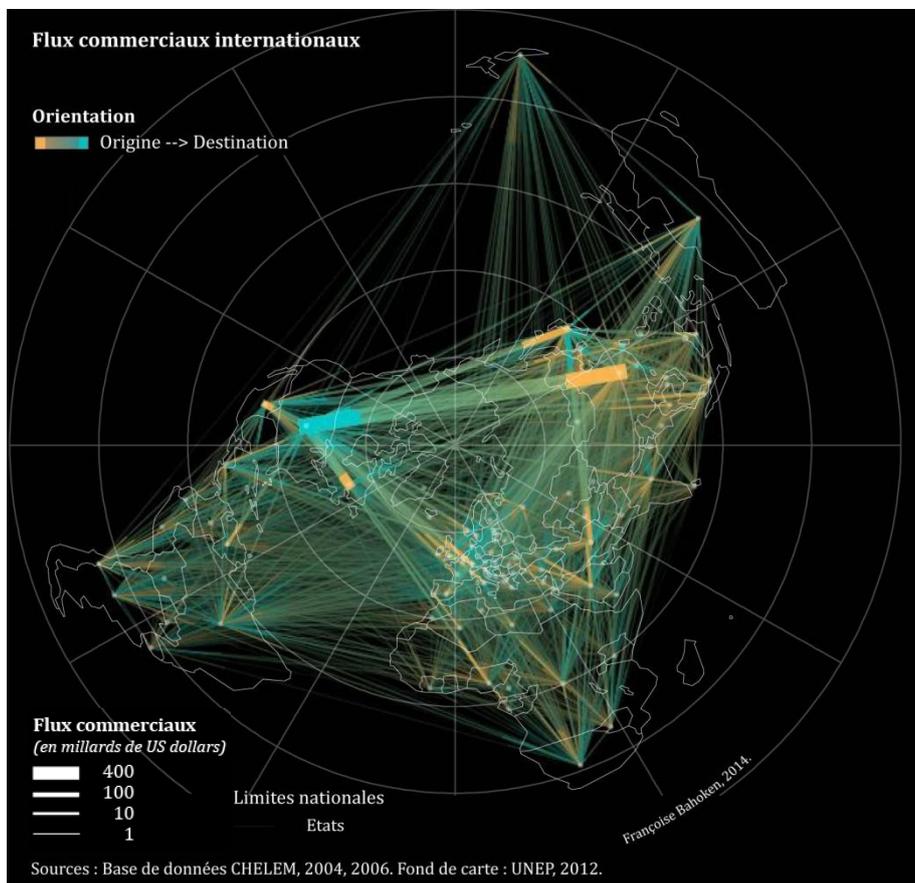
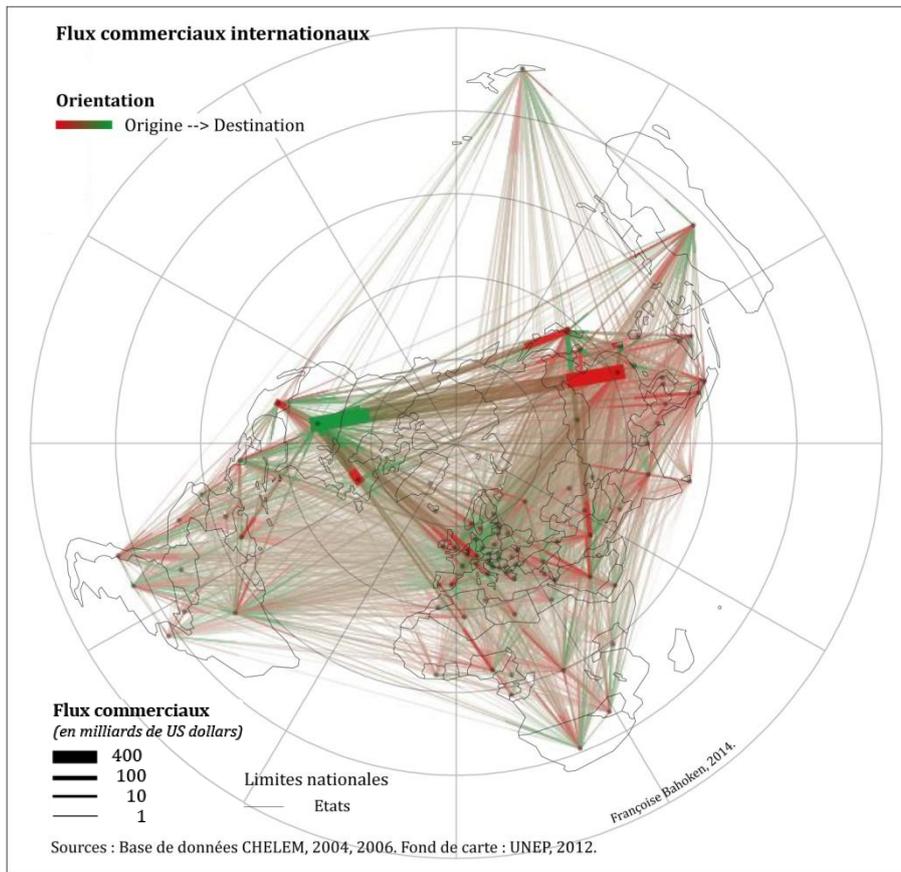
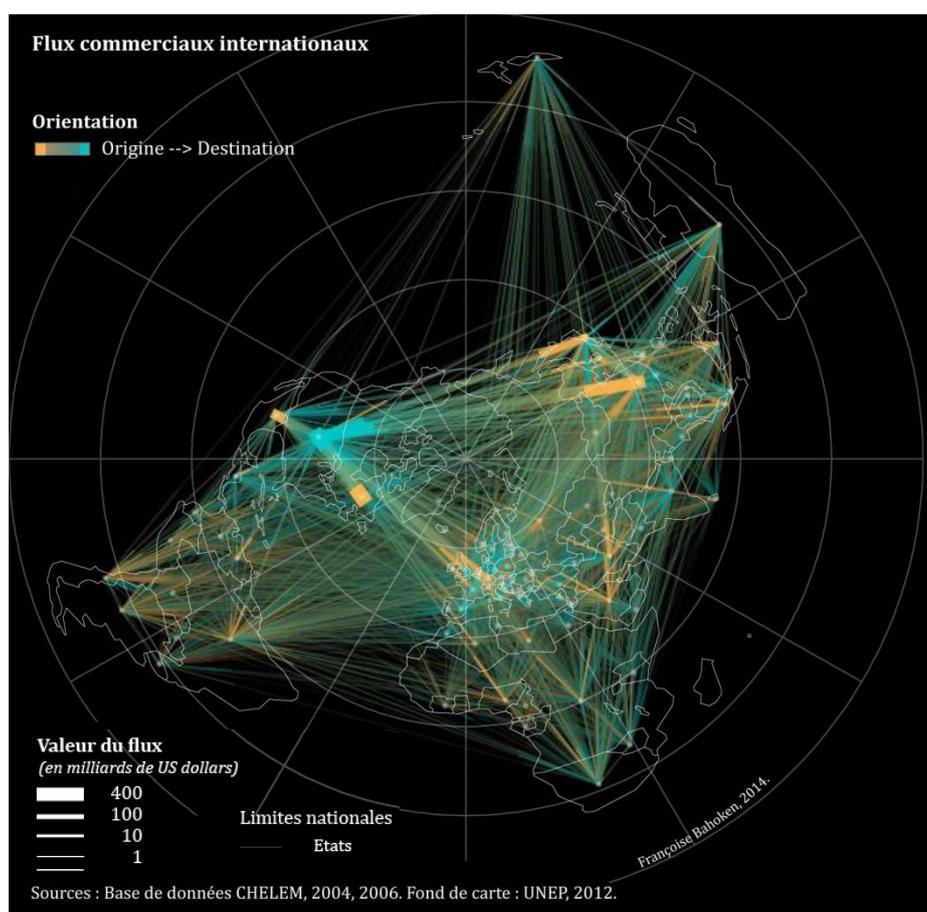


Figure 5-19. Cartographie du solde bilatéral (Fij) des échanges commerciaux mondiaux



La cartographie de (Fij) est susceptible de réduire fortement l'information représentée, sur le plan thématique : un certain nombre de liens aura des valeurs de solde bilatéral faibles voire nulles, l'objectif des états étant avant tout d'équilibrer leur balance des paiements. De ce point de vue, il peut être opportun de représenter les flux d'exportation distinctement de ceux d'importation, à la place des transferts nets bilatéraux (Fij). La Figure 5-20 présente une carte des flux d'exportation (Fij) commerciaux mondiaux qui peut naturellement être également interprétée en termes de flux d'importation (Fji), selon que l'on s'intéresse aux teintes orangées (à l'émission) ou aux teintes bleutées (à la réception).

Figure 5-20. Cartographie des flux commerciaux mondiaux d'exportations (Fij) ou d'importation (Fji)

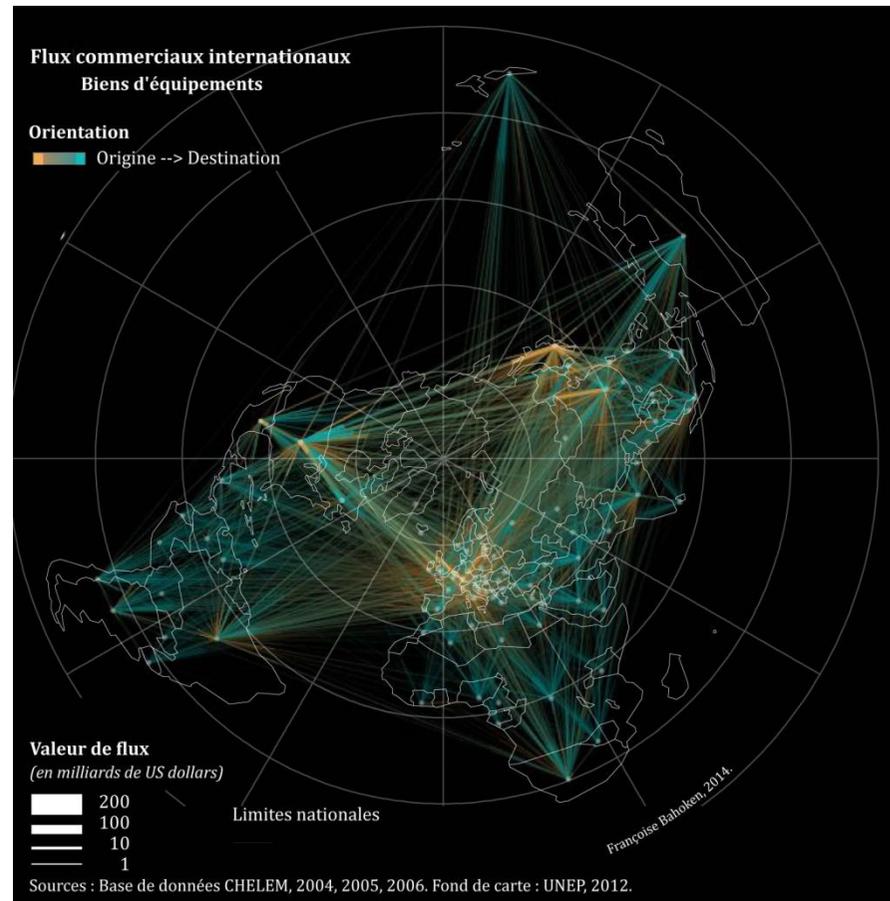
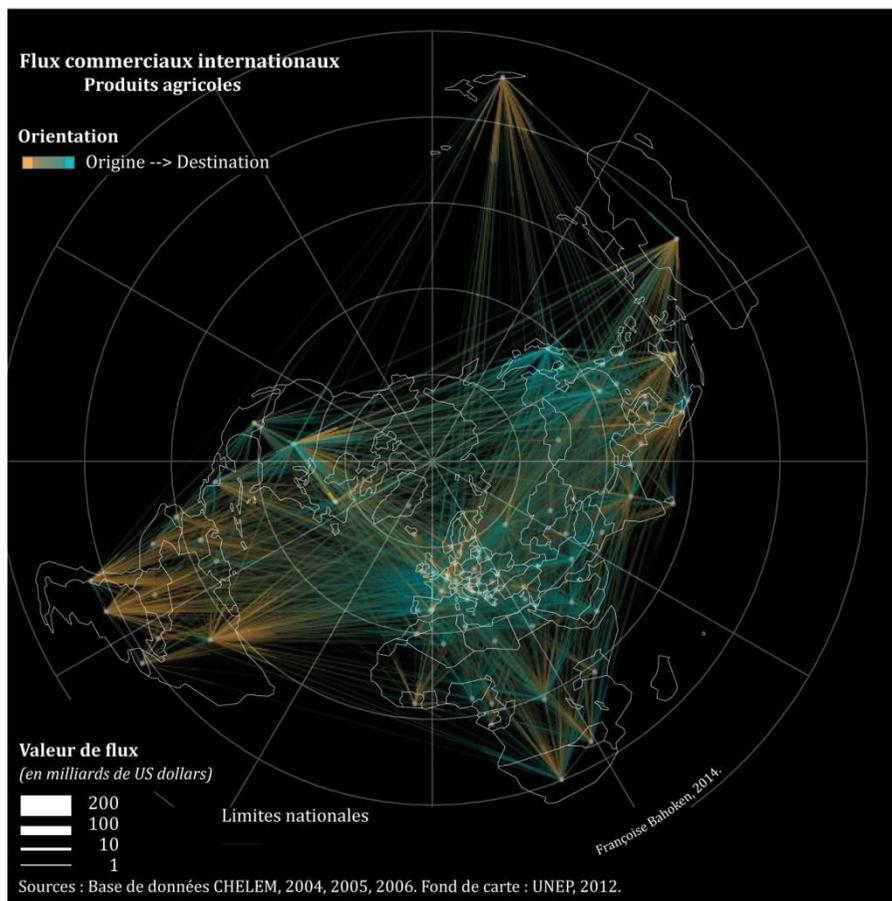


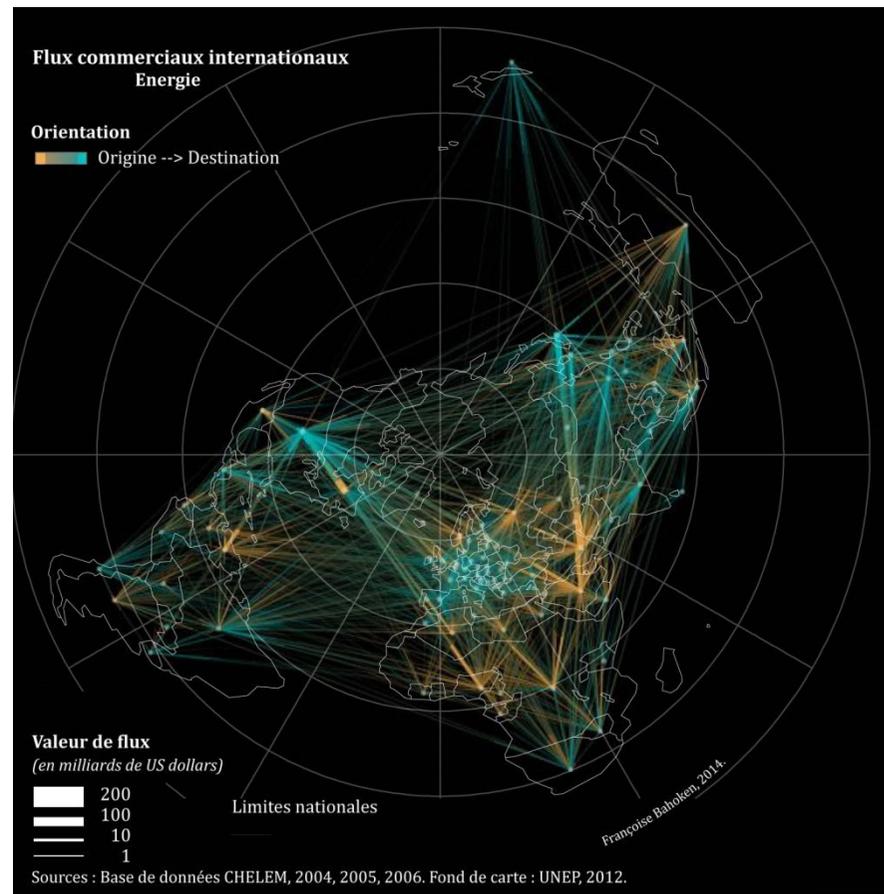
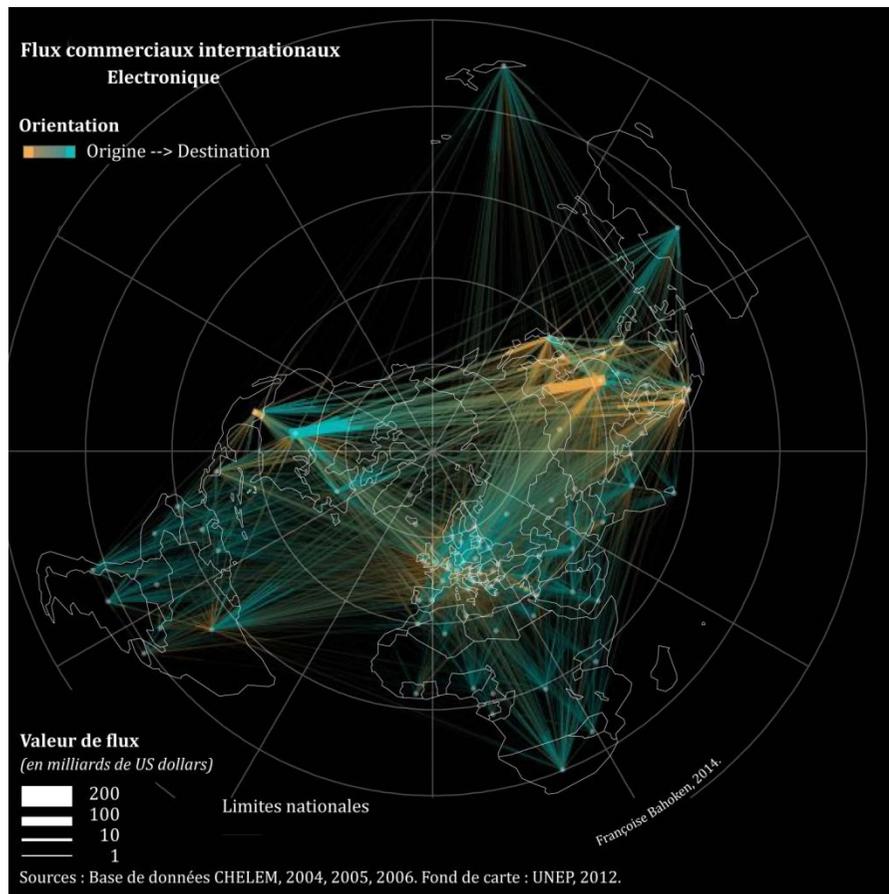
Si la carte des flux de marchandises mondiaux varie du point de vue du pays (exportateur *versus* importateur), elle est susceptible de varier également en fonction du type de marchandises, si la matrice est décomposable en catégories de produits différentes. On imagine aussi aisément la variation de la carte selon qu'elle s'applique à un produit particulier et aux flux d'importation ou d'exportation.

La Figure 5-21 présente une collection de quatre cartes de flux d'importation commerciaux, respectivement de produits agricoles (AGR), de biens d'équipements (EQU), d'électronique (ELE) et d'énergie (ENE). Ces cartes présentent une légende commune de manière à être directement comparables entre elles. Sur le plan thématique, ces cartes mettent en évidence la dissociation des lieux de production et de consommation à l'échelle mondiale.

Figure 5-21. Cartographie du solde bilatéral des échanges commerciaux (Fij)

Cartographie par catégorie de produits





Ces cartes permettent de visualiser la division internationale du travail ainsi que les relations commerciales selon une approche de type centre-périphérie, pour une sélection de produits. Leurs motifs généraux sont à comparer aux résultats sur l'évolution des trois macro-régions mondiales, par catégories de produits, obtenus par C. Grasland et G. Van Hamme (2010). Les auteurs font en effet référence au textile, à l'industrie automobile (et dans une moindre mesure à la chimie) dont ils analysent le « *poids total* », équivalent à (Fij^+) ; un « *indice d'asymétrie des échanges* » (tels que la matrice A_{ij} , définie dans la section 3.3.4, page 135) ; « *un indice de spécialisation des exportations* » ainsi que « *les flux origine-destination* » correspondant à (Fij) . Pour notre part, les cartes précédentes décrivent les soldes bilatéraux d'échanges commerciaux (Fij^-) des produits agricoles ; des biens d'équipements ; de l'électronique et du matériel électrique ainsi que de l'énergie.

Si la carte des transferts bilatéraux de produits agricoles apparaît peu contrastée en milliards de US\$ par rapport à toutes les autres catégories de produits sélectionnés, elle frappe par sa spécialisation. Sur le plan thématique, les échanges agricoles mettent en évidence l'importance des périphéries dans l'approvisionnement alimentaire des principaux centres mondiaux, mais aussi de leur voisinage. Elle peut être opposée aux cartes décrivant des flux de biens à haute valeur ajoutée. Ainsi, la carte des biens d'équipement est, à l'inverse de la précédente, focalisée sur les centres, en particulier sur l'Asie et les pays producteurs (exportateurs) de biens d'équipement présentant des valeurs monétaires élevées : elle est révélatrice de leur domination sur les pays consommateurs, qui correspondent au reste du monde.

Pour les flux d'énergie, l'importance de la péninsule arabique, des pays du golfe et des états situés dans le voisinage européen, en particulier la Russie, est indéniable dans l'approvisionnement des pays d'Europe occidentale et d'Amérique du Nord. La carte des flux d'électricité décrit, quant à elle, les flux les plus onéreux (la *largeur* des figurés y est beaucoup plus importante que dans les cartes précédentes), majoritairement émis par les pays d'Asie du sud-est. Sur la carte des biens d'équipement, on voit bien la surreprésentation des trois pays asiatiques (Malaisie, Philippines, Thaïlande), « [...] *autrefois spécialisés dans l'exportation de produits primaires, les tigres asiatiques sont maintenant positionnés sur les biens intermédiaires et les biens d'équipement, et dans une moindre mesure sur les biens de consommation* » (De Saint Vulry, 2008 : 17). Enfin, la carte des flux en biens d'équipements est la seule de cet ensemble où le rôle des pays européens exportateurs à cette échelle mondiale est visible.

On montre ainsi que les solutions proposées sont générales et généralisables à tous les types de flux, elles sont applicables quelle que soit l'échelle de la carte. La *Taille*, la *Couleur* et la *Transparence* sont appliquées indépendamment d'autres considérations, sur les deux exemples thématiques présentés dans ce chapitre : la variation de la carte étant naturellement imputable à la zone d'étude concernée et/ou à l'échelle de la carte (département de l'Isère *versus* monde entier) ; sur le plan analytique à la thématique du flux (migrations *versus* total de marchandises ou catégories) ainsi qu'à la composante du flux représentée (l'ensemble, les flux émis ou reçus, le volume ou le solde bilatéral, ...).

Ce chapitre permet enfin de montrer les répercussions sur le plan thématique d'une amélioration obtenue au niveau méthodologique, en l'occurrence à un niveau exclusivement sémiologique.

Conclusion du chapitre

Les solutions liées aux aspects esthétiques et graphiques proposées dans ce chapitre sont directement liées aux nouvelles possibilités techniques de visualisation de l'information développées dans le champ de l'informatique graphique. Transférées à la cartographie des matrices de flux, elles ouvrent la voie vers des perspectives intéressantes de renouvellement de la carte des flux, tout en réduisant la *surcharge graphique* spécifique à l'*effet-spaghetti* révélant ainsi un *effet de profondeur des flux*. Les propositions portent tant sur le plan méthodologique, que cartographique et sémiologique (esthétique), articulant famille de matrice et méthode de cartographie.

La pratique classique de cartographie des matrices de flux, qui nécessitait jusqu'ici l'application *a minima* d'un critère de sélection ou de réduction de l'information à représenter, voit ses possibilités étendues grâce au renouvellement de la sémiologie de la *Couleur*. Deux directions ont été privilégiées: la première porte sur la modification morphologique de la teinte et la seconde est relative à la mise en œuvre du principe de l'opposition colorée au niveau des liens.

La possibilité technique de recourir à la *Transparence* pour cartographier des matrices de flux denses, permet de représenter toute l'information disponible dans la matrice et cela, sans souffrir de l'*effet-spaghetti*. En effet, elle autorise la mise en œuvre d'un *effet de profondeur* visuelle, autrement dit la perception des figurés les plus petits en valeur, qui sont placés à l'arrière-plan de la figure, en transparence de ceux qui sont situés au premier plan. Cette masse de petits figurés est à l'origine de la complexité graphique caractéristique de l'*effet-spaghetti*. Le fait que les plus longs d'entre eux présentent les teintes les plus transparentes – qui sont (ou peuvent être) perçues comme les plus claires – n'est pas gênant puisque leur *longueur* perçue est suffisante pour permettre leur discrimination visuelle, par rapport aux courts flux placés au premier plan. La perception des plus petits flux n'est pas amoindrie en raison d'un effet positif de ce qui est, dans la version classique de la carte, un problème lié à leur *couverture spatiale*. L'usage des bandes proportionnelles permet aussi de résoudre le problème lié à la représentation conjointe de l'effet des flux au niveau des liens et des lieux, ce qui a pour avantage d'annuler la contrainte du respect de l'*effet convergence* des flux vers les lieux les plus attractifs.

Enfin, cette nouvelle sémiologie graphique permet de mettre en évidence l'existence d'une hiérarchie des flux qui nous conduit à émettre deux hypothèses: d'une part, **l'existence d'une forme d'attractivité des lieux représentée ici par l'effet de polarité des bandes orientées, en opposition colorée pour traduire leur sens et cela, sans recourir à la flèche**. D'autre part, **l'existence de différentes classes de longueurs de bandes dirigées vers ces lieux attractifs ce qui, dans la réalité, correspond à différentes classes de distances parcourue par les flux**.

La question de la perception de ces classes apparentes de distances est révélée par une action au niveau sémiologique de la carte de flux, par la manifestation d'un *effet de profondeur*. Elle peut être validée sur le plan mathématique, c'est-à-dire autrement que par le truchement d'une variable visuelle. Par exemple, en recourant à une sélection des flux de manière à ne représenter que ceux qui parcourent (plus ou moins) une distance donnée. Ces considérations vont faire l'objet du chapitre 7.

Avant cela, il nous semble important d'examiner, dans le chapitre suivant, la signification exacte de cette hypothèse **de perception de différentes classes de longueurs de bandes correspondant à des classes de distances parcourues**, car elle révèle en réalité un paradoxe. En effet, nous avons vu, jusqu'ici, que la cartographie de flux consistait à appliquer la *Taille*, ce qui correspond à une action sur la seule dimension de *largeur* du figuré (et non sur sa *longueur*). On peut donc s'interroger sur la signification de la perception de cette dimension du figuré qui n'est pourtant jamais mobilisée dans le processus de cartographie de flux. Cette interrogation qui va se révéler fondamentale fait l'objet du chapitre 6.

6^{ième} CHAPITRE. D'une cartographie de flux (Fij) à une cartographie de mouvements (Mij) et inversement

« Carte des déplacements de loisirs autour de Bratislava et Trnava [...]. Bandes de largeurs, proportionnelles aux flux. Par P. Mariot, *Geographicky Casopis* 1976 n°1 ».

Roger Brunet (1987).

Introduction du chapitre

Dans ce sixième chapitre, nous revenons aux aspects théoriques qui concourent à la réalisation d'une carte de flux, en examinant les questions liés à la construction géométrique du figuré (la flèche ou la bande) et ceux liés à la compréhension de l'information représentée. L'objectif général est, d'une part, d'examiner de plus près le paradoxe que nous avons identifié à la fin du cinquième chapitre, selon lequel l'information qui est représentée sur une carte de flux diffère de celle qui est perçue. D'autre part, de définir les deux cartes de flux commerciaux méditerranéens présentées par J. Bertin, sur la Figure 4-17 (page 173).

Au-delà des aspects de sémiologie qui ont permis d'ouvrir la voie vers la cartographie de *toute l'information disponible* dans une matrice de flux, l'un des résultats montre la possibilité de visualiser la masse de petits flux diffus simultanément aux plus gros, qui sont moins fréquents mais représentent dans l'ensemble une valeur importante. Les petits flux correspondent aux figurés les plus fins, de même qu'aux plus longs qui traversent la carte de part et d'autre de manière anarchique, générant ainsi la partie de *l'effet spaghetti* correspondant à un degré de *couverture spatiale* trop important.

La possibilité de percevoir plusieurs classes de *longueurs* de figurés en même temps que des variations de leur *largeur* est rendue possible grâce au recours à la *Transparence*. Cette possibilité n'est pas aussi anodine qu'elle en a l'air. Elle correspond sur le plan thématique à la problématique récurrente de la courte et de la longue distance, perçues par les différences de *longueur* des figurés. Cela pose la question d'un arbitrage sur la pertinence d'une cartographie des figurés les plus courts, correspondant à des mobilités locales, ou des plus longs, décrivant des flux d'envergure nationale ou internationale. Cette question sera examinée dans le chapitre suivant, aussi convient-il d'examiner d'abord la signification du « *comment* » la carte de flux est réalisée, en mobilisant la *largeur* du figuré, conjointement ou non avec sa *longueur* puisque l'œil perçoit les variations de ces deux dimensions simultanément. C'est cette possibilité qui pose hypothèse **d'un décalage entre l'information de flux disponible dans la matrice (Fij) et représentée sur une carte, et la perception qui en ressort, dont nous verrons qu'elle est similaire à un mouvement¹⁷⁵ (Mij).**

¹⁷⁵ La notion de mouvement utilisée ici, notamment en référence au cadre d'analyse des flux migratoires, ne doit pas être confondue avec celle de mouvement relatif au transport des biens. La mesure de ces deux mouvements, et leur implantation spatiale, diffèrent fondamentalement. Dans le second cas, le mouvement correspond à l'émission et/ou à la réception d'un ou de plusieurs colis en un point donné, à l'aide d'un véhicule motorisé (Gerardin et al., 2008). Il fait

La pratique usuelle de la cartographie de flux est-elle réellement pertinente ? Est-il si évident que cela de réaliser des cartes de flux en ne mobilisant qu'une seule dimension du figuré (sa largeur), alors que l'œil humain en perçoit plusieurs ? Et notamment la surface ?

Ce paradoxe fait naître un doute raisonnable quant à la validité de cette pratique. Nous proposons de l'examiner en nous intéressant aux aspects cognitifs de la cartographie de flux, c'est-à-dire à la manière dont les flux cartographiés sont perçus par un observateur et, *a fortiori*, celle dont ils sont tracés. En effet, ce paradoxe peut être levé en raisonnant à la manière de J. Bertin (1967) et, plus généralement, selon une approche sémiotique. Les travaux de cet auteur ont porté pour une grande part sur les aspects cognitifs des signes graphiques, c'est-à-dire sur la manière dont le cerveau humain comprenait les informations perçues.

Nous allons ainsi démontrer que ce paradoxe, tout en étant bien réel, n'est qu'apparent, dans la mesure où il ne conduit pas à la réalisation d'une carte erronée. Il correspond à la solution classique dont nous démontrerons la légitimité, en examinant les aspects sémiotiques de la carte de flux. La réflexion sera menée aux niveaux cartographiques et sémantiques de la réalisation cartographique, considérant le langage cartographique comme un mode d'expression ayant des répercussions au niveau cognitif de la perception du phénomène. La réflexion cartographique porte sur la construction du figuré qui décrit le flux ; elle correspond au *signifiant* tandis que la réflexion sémantique concerne le *signifié*, c'est-à-dire la manière dont l'information cartographique (le figuré) est perçue par le mode de représentation adopté (le choix du *signifiant*).

La Figure 6-1 située page suivante présente le cadre théorique de notre réflexion. Nous l'illustrons en prenant l'exemple de la cartographie du (Fij⁺) réalisée, comme nous l'avons vu dans le chapitre 3, à l'aide d'une bande. Nous examinerons ainsi une série de quatre hypothèses successives correspondant aux différentes cases de cette figure. L'exercice n'étant pas aisé, nous allons nous référer régulièrement à cette Figure 6-1 que nous expliciterons au fur et à mesure de notre démonstration, devant nous conduire à valider ou non, tour à tour, les quatre hypothèses mentionnées dans chacune des cases.

Dans une *logique de flux*, telle que celle qui prévaut depuis le début de cette thèse, le tracé du figuré est généralement¹⁷⁶ le résultat d'une action menée sur sa *largeur*, en application de la *Taille*, voire sur sa *surface*, en application de la *Valeur*, par exemple. Sauf dans les travaux de W. Tobler (1969), de J.-C. Müller (1979) et de A. L'Hostis (1996, 2003, 2008) déjà mentionnés et sur lesquels nous reviendrons¹⁷⁷ dans la seconde section de ce chapitre.

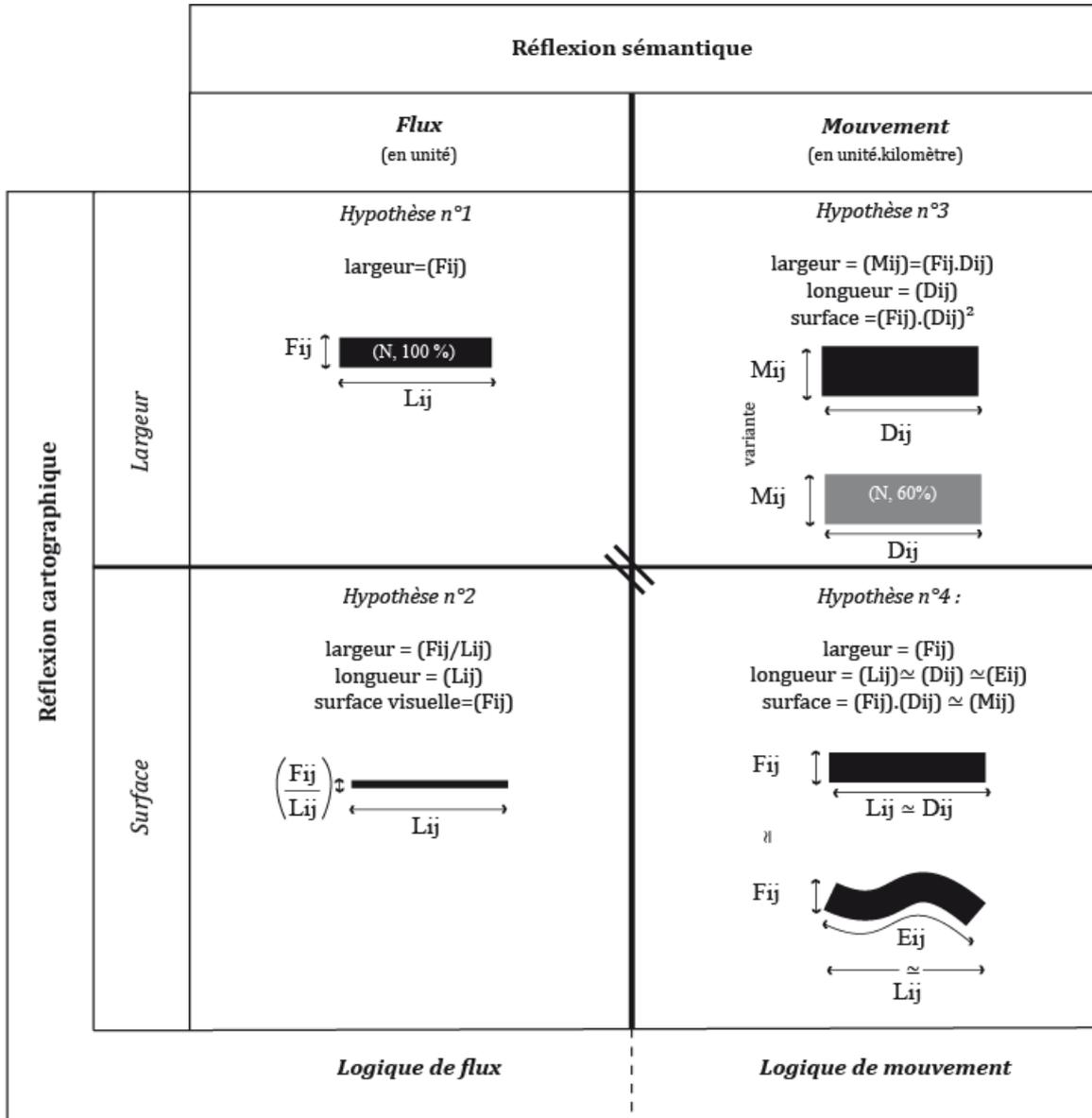
Autrement dit, **la réalisation d'une carte de flux ne mobilise que la largeur du figuré et non sa longueur, elle n'implique que la valeur (Fij) disponible dans la matrice sans se préoccuper de l'espace, des positions (x, y) des lieux. Cette pratique, aussi étrange qu'elle puisse apparaître au géographe, lorsqu'elle est ainsi formulée, est caractéristique de l'approche usuelle de la cartographie de flux (voir chapitre 3).** Elle correspond à la première hypothèse que nous allons traiter, intitulée Hypothèse n°1 sur la Figure 6-1.

alors référence à une opération ponctuelle (de livraison et/ou d'enlèvement de marchandises) et l'analyse est réalisée du point de vue des lieux et non des liens. Dans notre cadre d'analyse, le mouvement est bien appréhendé en première instance comme un flux, du point de vue des liens valués.

¹⁷⁶ Nous avons vu (chapitre 4) que certains auteurs mobilisaient la *hauteur* du figuré, lorsque la carte est réalisée en trois dimensions : il s'agit d'un élément de détail qui consiste à appliquer la *Taille* non seulement sur la *largeur* du figuré mais également sur sa *hauteur*. Cette proposition n'est pas de nature à modifier le raisonnement que nous menons ici, elle correspond plutôt à une double application de la *Taille*.

¹⁷⁷ Ces travaux portent sur la cartographie de flux affectés sur réseaux (planaires) et, s'ils étaient appliqués aux flux (réseaux non planaires), ils conduiraient à des cartes que l'on ne saurait considérer comme classiques, elles mobilisent des solutions liées aux transformations que nous ne mobilisons pas dans cette thèse.

Figure 6-1. Cadre sémiotique de réflexion sur la construction de la carte de flux : aspects cartographiques et sémantiques



Dimensions du figuré linéaire

Largeur

Fij : quantité transférée d'un lieu à un autre, dans un plan
Mij : quantité transférée d'un lieu à un autre pondérée par la distance parcourue

Longueur :

Lij : distance cartographique perçue (= longueur du figuré)
Dij : distance géographique, euclidienne ou réseau (= longueur du figuré)
Eij : distance empirique ou coût généralisé de transport (= longueur du figuré)

Surface :

surface visuelle (N) : quantité de teinte noire, exprimée en pourcentage (%) de CMJN
surface géométrique (S) : produit de la longueur par la largeur du figuré

Françoise Bahoken, 2015.

Dans la première section, nous examinerons une série de quatre hypothèses successives correspondant aux différentes cases de la Figure 6-1. L'objectif est d'étudier le processus de cartographie de flux d'un double point de vue : cartographique et sémantique, du *signifiant* et du *signifié*.

Nous analysons l'hypothèse n°1 selon laquelle **l'œil ne percevrait que les variations de largeur du figuré ce qui constitue, d'après nous, un paradoxe**. Le figuré qui illustre le flux est une forme géométrique dont la construction mobilise deux dimensions. Ce constat suggère **la possibilité de percevoir la surface visuelle du figuré, représentée par la quantité de teinte noire apparente (N, 100%)** qui correspond à une seconde hypothèse. Si elle est plausible, sa mise en œuvre pratique s'avèrera intenable en l'état, pour des raisons d'ordre méthodologique et thématique. La notion de *surface visuelle* nous conduit toutefois vers celle de *surface géométrique*. Cette nouvelle dimension mobilise la *longueur* du figuré qui correspond en réalité à la *distance cartographique*. L'examen de cette troisième hypothèse, selon laquelle **l'œil percevrait des variations de la surface des figurés, nous amène à introduire le rôle joué par l'espace au travers de l'implantation spatiale des lieux et de leur espacement**, notamment avec le recours à la notion de *distance parcourue* par le flux. Nous présentons ainsi, dans la seconde section, les trois formes de distances qui sont mobilisables en cartographie de flux à savoir : la *distance cartographique* (Lij), la *distance géographique* (Dij) approximée par la *distance euclidienne* et la *distance empirique* (Eij). Cette discussion sur les distances marque le passage d'un raisonnement en termes de *logique de flux* (Fij) à un raisonnement fondé sur une *logique de mouvement* (Mij), en inscrivant ainsi notre réflexion dans un espace géographique, dans un territoire alors qu'il était question d'un espace (ou d'un plan) quelconque. Dans la troisième section, nous définissons la quantité de mouvement de (Mij) au regard de celle de flux (Fij) puis examinons l'hypothèse n°3 (voir Figure 6-1). Elle nous conduit naturellement à l'hypothèse n°4 qui **inscrit la distance parcourue comme composante intrinsèque de la perception des flux et cela, quelle que soit la quantité dont il est question : un flux nécessairement spatialisé ou un mouvement** dont on connaît inévitablement la trajectoire spatio-temporelle. Elle nous conduit à la conclusion selon laquelle les hypothèses n°1 et n°4 sont **identiques au niveau cartographique de leur construction, mais différentes au niveau sémantique ce qui prouve qu'en réalité, la cartographie de flux est une représentation approximative du mouvement**.

6.1 La réalisation d'une carte de flux selon une *logique de flux* (Fij)

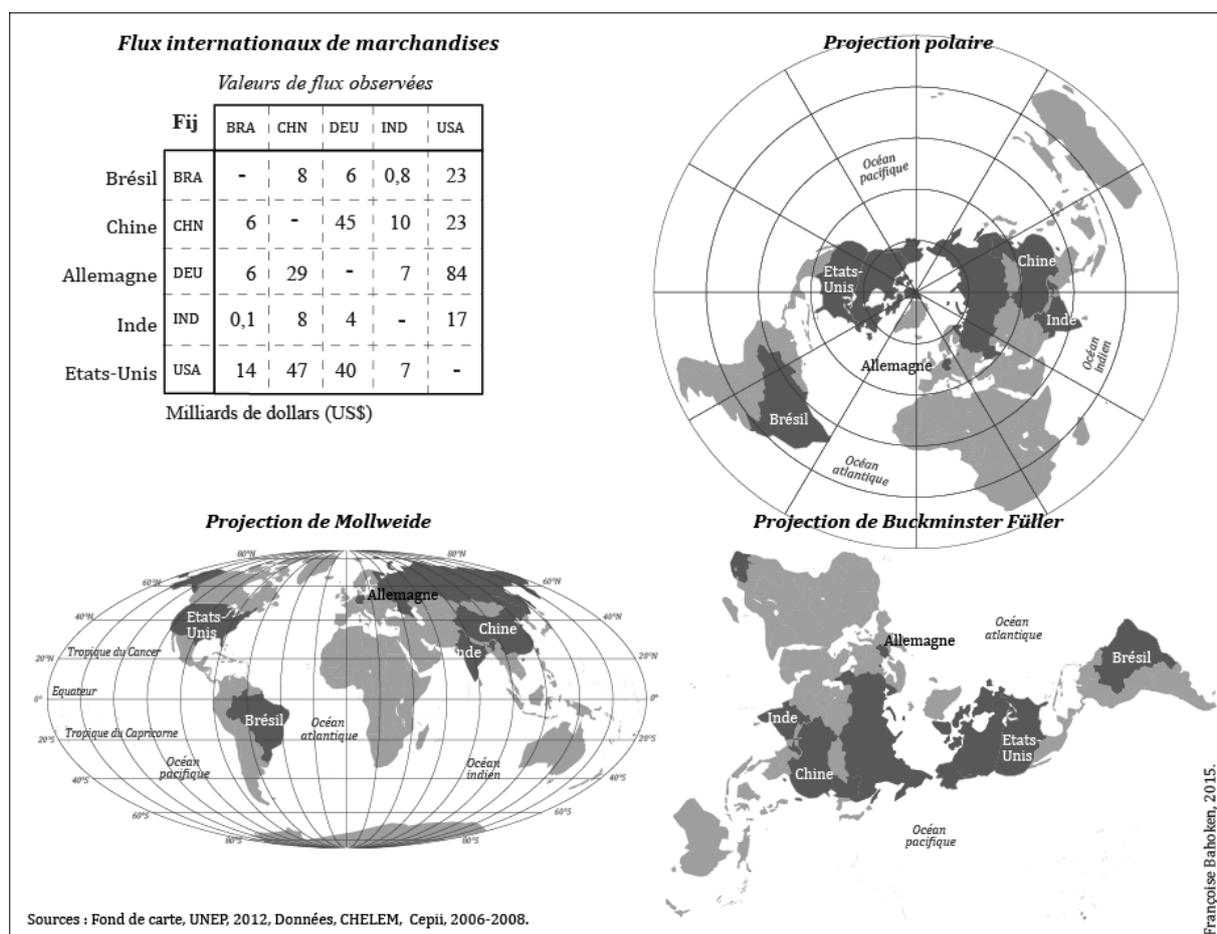
L'observation de la manière dont les cartes de flux sont construites et perçues par un observateur statique nous conduit à revenir sur l'approche classique de la cartographie de flux, telle que nous l'avons présentée dans le chapitre 3. Cette approche correspond à la première des hypothèses suggérées sur la Figure 6-1 (page 249). Elle émane du paradoxe que nous décrivons ci-dessous à partir d'un exemple de cartographie d'une matrice de flux internationaux à l'échelle mondiale.

Nous mobilisons la même matrice de flux internationaux de marchandises issue de la base de données CHELEM (2006-2008) du CEPII, que dans le chapitre 5, moyennant une sélection de l'information. Les données (Fij⁺) concernent une sélection de cinq pays: les Etats-Unis (USA), le Brésil (BRA), l'Allemagne (DEU), la Chine (CHN) et l'Inde (IND). Ces données sont présentées sur la Figure 6-2, les pays sélectionnés étant localisés sur trois fonds planisphère de projection différente.

Les valeurs de la matrice de flux correspondent au total de marchandises échangées entre les cinq pays, il est exprimé en dollars américains (US\$). La projection cartographique de Mollweide-Babinet correspond à la vision du monde fréquemment utilisée en alternative à celle de Mercator considérée comme inappropriée à la cartographie statistique parce qu'elle conserve les angles et distord les surfaces. La projection de K. B. Mollweide (qui correspond à la projection holographique de J. Babinet) est équivalente et cylindrique. Elle respecte les surfaces, au détriment des angles, ce qui permet de visualiser, « *dans un rapport de similitude convenable* » (Joly, 1976 : 1), la répartition

globale des masses continentales (ou océaniques) sur la portion d'espace situé entre l'équateur et 40° à 45° de Latitude Nord et de Latitude Sud environ ; au-delà, les surfaces et les angles sont distordus d'où l'extension visible des pôles nord et sud. La projection de R. Buckminster Fuller (1946) est considérée comme universelle, car elle ne privilégie aucun espace sur le plan géométrique, géographique ou symbolique. En étant conforme, la vision du monde qu'elle propose ne présente pas de nord, ni de continuité territoriale, ce qui est déroutant. Si ce mode de construction déformant les orientations cardinales habituelles la rend difficile à lire, elle n'est pas inadaptée à la représentation de flux internationaux, comme l'illustre la Figure 6-3. La troisième projection utilisée est azimutale et équidistante, elle est centrée sur le pôle nord (polaire) et correspond à une vision du monde supposée correcte pour la représentation des flux.

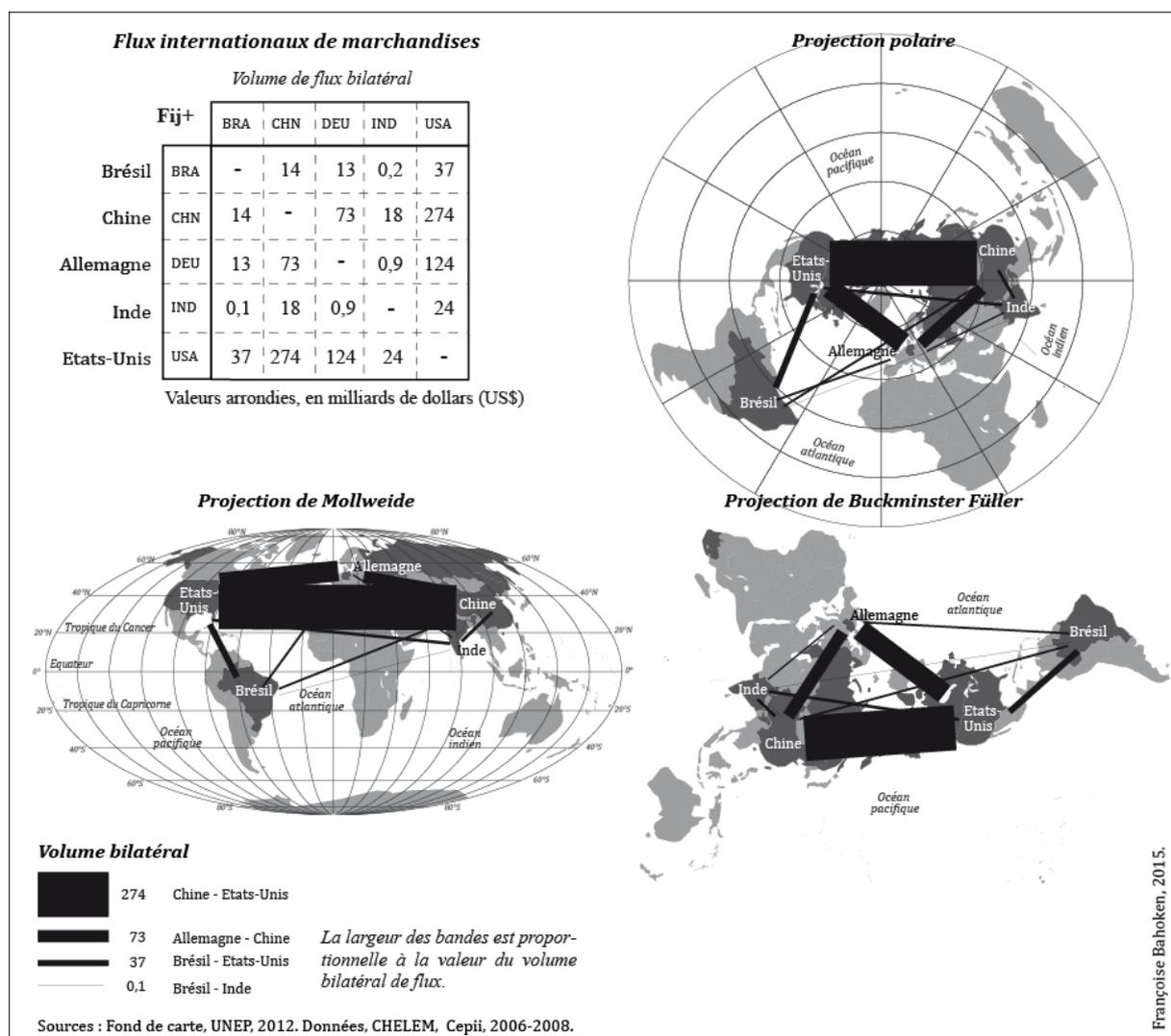
Figure 6-2. Présentation des données de l'exemple des flux mondiaux



Le choix d'un exemple pris à l'échelle mondiale, selon différents systèmes de projection cartographique, tient au besoin de la démonstration de mettre en évidence la variation de la perception de l'information de flux cartographiée à une même échelle, en fonction d'un changement de la position spatiale des lieux. D'après C. Grasland, elle sert aussi à démontrer par la cartographie que l'espace des positions (donc des lieux) ne peut être dissocié de celle des interactions (donc des flux) entre les lieux. En effet, « [...] nous pouvons trouver une démonstration claire de l'intérêt d'une approche symétrique des positions et des interactions à travers la notion de distance, ce qui constitue d'ailleurs la majeure contribution de W. R. Tobler [1969, 1970] à la géographie théorique et quantitative » (Grasland, 2009 : 3), nous y reviendrons dans la section 6.2.

La Figure 6-3 présente la cartographie du volume de flux bilatéral (F_{ij}^+) issu de la matrice de flux de la Figure 6-2, sur les trois fonds de carte sélectionnés.

Figure 6-3. Cartographie des flux internationaux (volume bilatéral) de marchandises

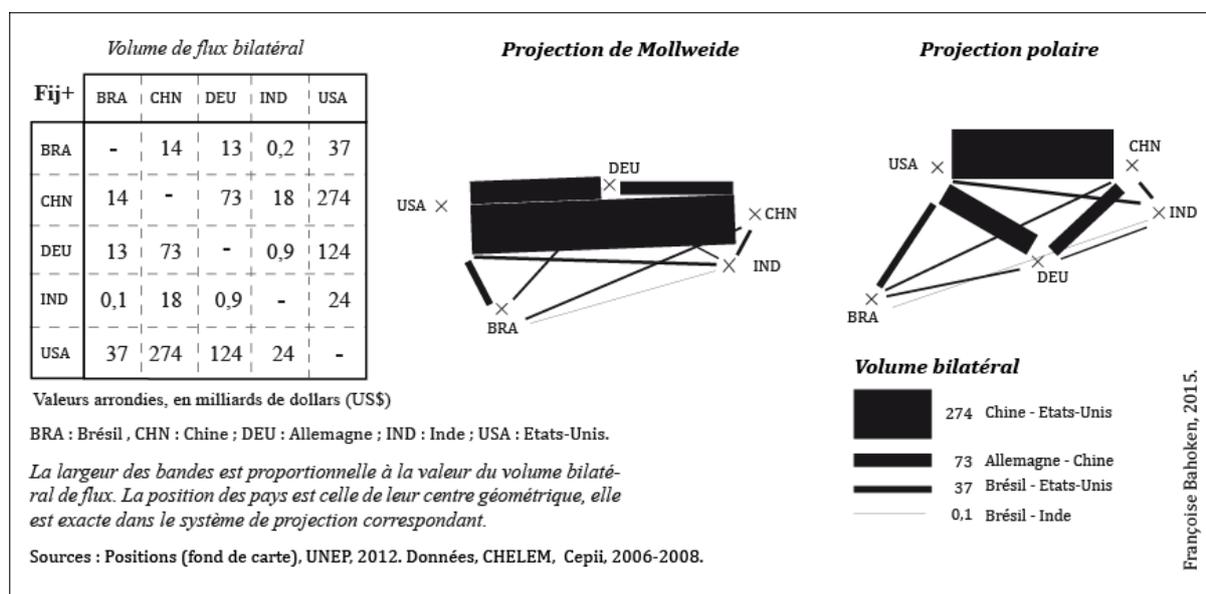


Afin de mieux percevoir les motifs de ces flux mondiaux, nous supprimons les fonds de carte et ne retenons que les tracés des flux projetés en Mollweide et polaire sur la Figure 6-4, le résultat obtenu avec la projection de Buckminster Fuller étant identique à la polaire.

D'après la Figure 6-4, il est intéressant de noter que **le changement de projection cartographique entraîne un changement de la position des lieux qui se traduit par une modification du motif des flux**. Dit autrement, un changement de projection entraîne une modification de l'information sur la structure spatiale des flux mondiaux. On observe en effet que dans le cas de la projection Mollweide, la structure spatiale de ces flux correspond globalement à un gros trait horizontal (USA-CHN), d'axe est-ouest qui est caractéristique d'un *effet d'alignement topologique* lié à la projection utilisée. La particularité de cet effet est d'empêcher la perception du phénomène – ici, en l'occurrence un processus de mondialisation des flux internationaux de marchandises –, ce qui traduit le caractère inadapté de ce type de projection. C'est pourquoi la réalisation d'une carte de flux internationaux à l'échelle mondiale nécessite le recours à un autre système de projection qui annihile cet *effet*

d'alignement topologique : la projection polaire (tout comme celle de Buckminster Fuller) font partie de celles qui conviennent.

Figure 6-4. Illustration du volume bilatéral (Fij⁺) de flux internationaux de marchandises



Cependant, si le passage à une projection polaire permet de corriger cet *effet d'alignement topologique*, il conduit dans le même temps à une transformation de l'information, à la perception d'une structure spatiale différente de la précédente, alors qu'il s'agit de la cartographie des mêmes flux réalisée à la même échelle. Sur la projection polaire, le motif de ces flux internationaux de marchandises correspond à un triangle (USA-CHN-DEU) dont la base serait est-ouest et non plus à un rectangle horizontal (USA-CHN), comme sur la projection de Mollweide. Ce résultat n'est pas spécifique à l'échelle mondiale, il est possible de l'observer à l'échelle locale, lors d'un changement de la position des lieux, ce qui prouve qu'il ne dépend pas de l'échelle spatiale. C'est donc **le changement de projection (de la position des lieux) qui impacte le motif des flux, quelle que soit l'échelle géographique d'observation. La raison tient au fait que sur une carte à petite échelle, outre la métrique utilisée, la valeur de l'espacement des lieux est dépendante du choix de la projection.**

Ce changement de l'information perçue de flux, en fonction de la position des lieux est-il normal ? N'y aurait-il pas un problème de perception du flux, lorsqu'il est cartographié de manière classique ?

Il y a un paradoxe qui correspond au décalage entre l'information cartographiée (Fij) et sa perception, dont on perçoit bien la variation de la quantité de noir en fonction de l'implantation spatiale des lieux. Sur la Figure 6-4, nous avons l'impression qu'il y a plus d'information sur l'axe USA-CHN de la carte en projection Mollweide, que sur le même axe en projection polaire. La raison tiendrait au fait que l'œil percevrait en réalité une variation de la quantité de teinte qui correspond à la surface géométrique des figurés qui illustrent les flux et non à leur seule *largeur*.

Pour examiner de plus près ce paradoxe, repartons de l'approche classique de cartographie de flux qui correspond à l'hypothèse n°1, dans la sous-section suivante.

6.1.1 L'hypothèse n°1 d'une perception des variations de 'largeur' du figuré : un paradoxe

La première hypothèse correspond à l'approche classique, à l'origine de notre étonnement.

On admet, en effet, généralement et sans discussion, le fait que ce soit la *largeur* (que nous substituons au terme d'épaisseur¹⁷⁸) du figuré qui représente – de manière proportionnelle grâce à l'application de la *Taille* – la valeur du flux (Fij). D'ailleurs, les outils¹⁷⁹ de cartographie automatique des flux sont tous fondés sur ce principe. Raisonner ainsi c'est supposer que l'œil ne perçoit que les variations de *largeur* de la bande et non celles de ses autres dimensions. Est-ce vraiment le cas lorsque l'on observe la Figure 6-3 ? Qu'en est-il de la carte de P. Mariot (Figure I-1-1, page 13) ?

Supposer cette première hypothèse exacte présente deux conséquences théoriques fortes : cela signifie d'une part, que **l'œil ne perçoit qu'une seule dimension géométrique** et d'autre part, que **l'espace n'est pas pris en compte dans le processus de cartographie de flux**. Pour démontrer cette affirmation, il convient de revenir sur la définition du concept de *flux* et la manière dont on le représente sur une carte.

Définition du flux : Le flux correspond à une quantité mesurable d'objets mobiles, à un stock ou à un effectif, qui a changé de zone (ou de portions d'espaces délimitées ou représentées par un point).

Cette définition stricte du flux ignore le mode de déplacement de la quantité dont il est question, par conséquent son caractère matériel ou immatériel. La définition du flux ne prend pas en compte le cheminement dans l'espace, lequel n'interviendra donc pas dans le processus de cartographie de flux.

En effet, l'approche classique de la cartographie de flux, telle que nous l'avons présentée dans le chapitre 3, consiste à n'agir que sur la seule *largeur* du figuré, sans se préoccuper ni de ses autres dimensions ni de son implantation (x, y). Le cartographe trace un lien valué entre deux lieux qui sont placés dans un plan quelconque, sans se préoccuper de la métrique et de la *longueur*, pour pouvoir être perçus sous une forme (carto)graphique. La position des lieux est par conséquent arbitraire, elle peut être modifiée à tout moment pour satisfaire la visualisation aisée des relations dans un plan qui peut être circulaire ou encore sous la forme d'arbre, pour ne citer que ces deux exemples de placements dans un graphe. Le plan dont il est question correspond à n'importe quel espace, ce qui explique le fait qu'il ne soit jamais question de métrique dans le processus général de cartographie de flux. Le cartographe n'en fait jamais état, à notre connaissance.

C'est parce que la localisation des espaces mis en relation par un flux (et donc des distances, ...) n'intervient jamais dans le raisonnement cartographique que nous avons pu considérer, dans le chapitre 2, le graphe comme une solution légitime permettant de visualiser des flux. Après tout, la quantité qui se déplace entre deux portions de territoires ne nécessite pas que ces derniers soient localisés dans un espace géographique, leur position dans un plan, d'au moins une dimension, étant suffisante pour autoriser leur représentation graphique. L'espace n'ayant pas forcément d'intérêt dans la représentation graphique des flux, toutes les variables traditionnelles qui permettent de l'appréhender deviennent accessoires.

Ce constat du caractère secondaire de l'espace dans la définition du flux induit une conséquence forte sur le plan sémiologique, qui justifie à elle seule le principe théorique de sémiologie cartographique du flux. Nous en arrivons à la conclusion suivante. **La pratique classique de la cartographie de flux** (voir chapitre 3 et chapitre 4) **ne correspond pas à l'application *stricto sensu* du principe de**

¹⁷⁸ L'épaisseur d'une forme géométrique renvoie à une notion de volume que nous mobilisons déjà largement par ailleurs (voir chapitre 1 et chapitre 3), de même qu'à celle de troisième dimension que nous avons écarté de notre champ d'analyse. La *largeur* nous apparaît mieux correspondre à notre propos qui considère les différentes mesures d'un figuré géométrique, la bande, qui illustre un flux que nous représentons en deux dimensions. Ainsi, la *largeur* correspond à la mesure du plus petit côté de la bande, elle est nécessairement liée à celle de la *longueur* qui correspond à son plus grand côté. Elle apparaît ainsi mieux adaptée au caractère géométrique de la forme qui caractérise le flux.

¹⁷⁹ La deuxième section de la bibliographie présente une sélection d'outils permettant de cartographier des flux.

proportionnalité entre la valeur disponible dans la matrice (Fij) et la surface du figuré. Autrement dit, le figuré qui illustre la valeur (Fij) ne correspond pas à son rapport de surface représenté par le produit (longueur*largeur). De là découle le paradoxe.

Etant donné que cette première hypothèse correspond à la pratique usuelle de cartographie de flux, nous la qualifions de *logique de flux*. La particularité de la *logique de flux* est donc de privilégier la représentation de la valeur (Fij) du flux au détriment de celle des positions des lieux dans l'espace. Cette posture n'est pas absurde, elle est logique. Elle correspond d'ailleurs à la solution officielle suggérée par J. Bertin et reprise par les différents auteurs ayant examiné la question du tracé du figuré (Tobler, 1987 ; Dorling, 1991, Zanin, Trémolo, 2003). C'est aussi la logique que nous avons également appliquée jusqu'ici pour réaliser l'ensemble des cartes de flux présentées.

Dans une *logique de flux*, seule la valeur (Fij) importe dans la représentation (carto)graphique. Est-elle pour autant perçue comme la plus importante ?

D'après la Figure 6-4, la réponse est négative, puisque ce qui semble primer, c'est une variation de la quantité de teinte et non de sa seule *largeur*. Cette quantité de teinte plus ou moins importante correspond à la *surface visuelle* du figuré qui serait la dimension qui, parce qu'elle est instantanément perçue comme telle, symboliserait le flux.

Cela signifie que pour une carte de flux où la couleur et la Valeur des bandes représentant les flux seraient uniformes (e.g. 100% de noir comme sur la Figure 6-4), l'œil détecterait une masse visuelle égale non pas à la somme des valeurs de flux (comme nous l'avons indiquée dans le chapitre 3) mais plutôt à la somme des valeurs d'une autre quantité définie par la surface visuelle du figuré, à savoir un mouvement (le produit des flux par une distance).

Sur la Figure 6-4, si l'on observe le couple USA-CHN sur les deux projections, on observe que le tracé des flux qui est perçu *instantanément* est celui réalisé sous projection Mollweide : il s'agit bien du figuré qui présente la *surface visuelle* la plus importante, la valeur de la *largeur* étant identique dans les deux cas.

Nous en concluons que le raisonnement cartographique mené selon une *logique de flux* intègre la *perception du mouvement dans la représentation du flux*. La distance parcourue, en considérant qu'elle est donnée par le système de projection, est représentée par la *distance cartographique*, puisque l'information est spatialisée. Mais elle n'intègre pas la *distance géographique*, celle qui a été parcourue, à aucun moment de l'analyse, d'où le décalage entre l'information de flux (Fij) qui est représentée et la perception de sa représentation. Si l'on poursuivait l'analyse au niveau des lieux, nous arriverions à la conclusion que l'observateur perçoit au droit des cercles proportionnels, sur une carte de flux de type lien-lieux, une représentation de la somme des *surfaces* des bandes et non de leur seule *largeur*. Or, cela ne correspond pas au résultat que l'on souhaite *a priori* montrer dans une carte de flux.

Dans le cadre d'une *logique de flux*, l'utilisation classique de la *Taille* entraînerait comme un effet d'optique similaire à la mauvaise utilisation de la taille de symboles proportionnels, ne tenant pas compte de la surface. Il conduit à une asymétrie négative entre la perception de la quantité de mouvement dépensée et sa représentation.

Ce constat conduit à deux biais de lecture des cartes de flux : le premier concerne la perception erronée des quantités et des distances parcourues, à la source potentielle d'erreurs dans l'interprétation des flux cartographiés ; le second biais est lié à la difficulté de l'analyse des cartes du volume bilatéral de flux, qui ne décrivent pas correctement les motifs spatiaux. Pour les éviter, deux solutions sont envisageables. La première solution est d'ordre sémiologique, elle consiste à utiliser le procédé du triangle (ou tête de flèche) introduite par J. Bertin (voir Figure 4-31, page 194). Ce dernier permet, en effet, d'agir sur la *surface* du figuré (en mobilisant sa longueur, sa hauteur et sa surface) sans affecter

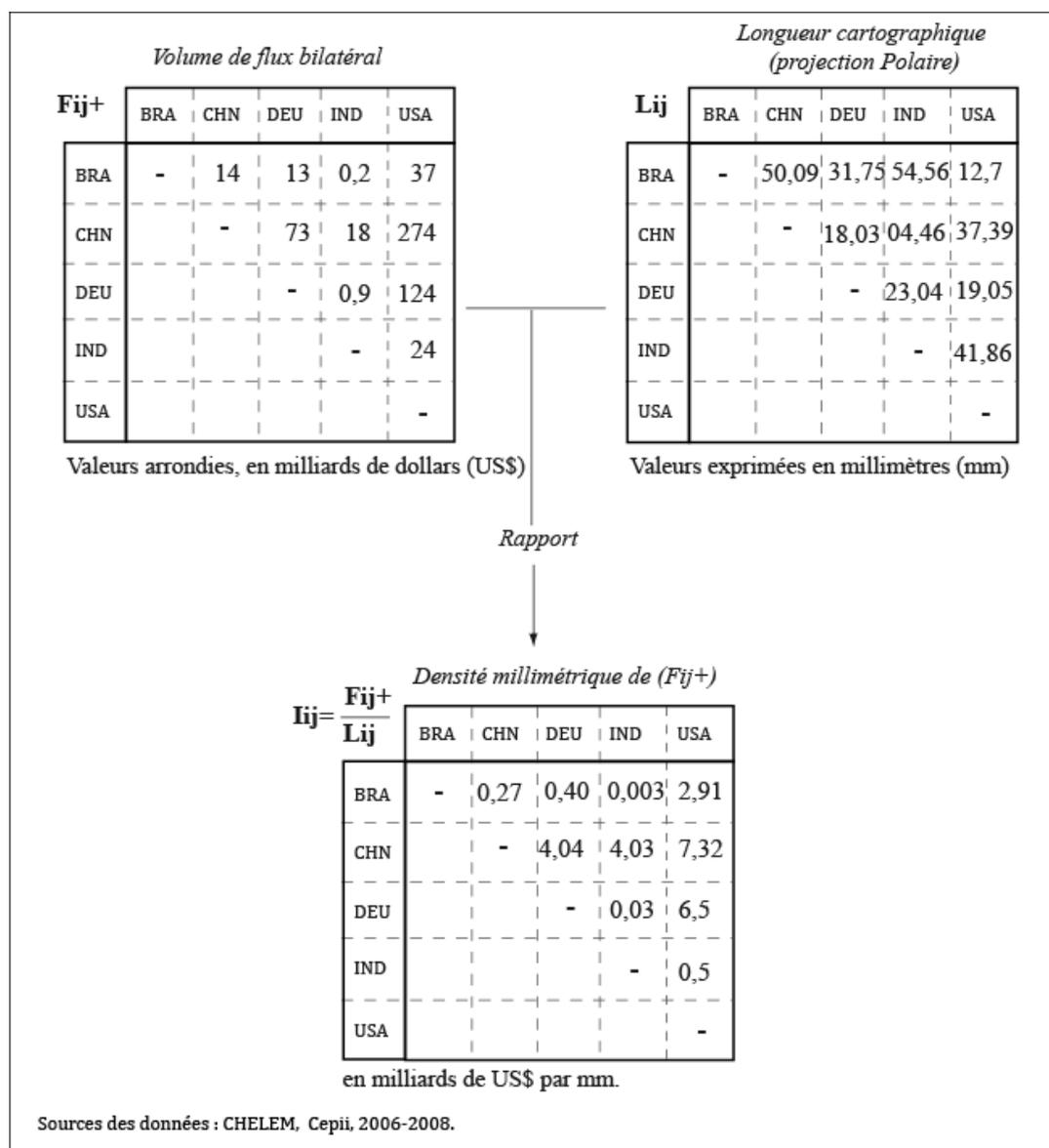
la *longueur* du corps ; la seconde solution consiste à corriger la perception des surfaces visuelles. Cette éventualité nous conduit vers une seconde hypothèse.

6.1.2 L'hypothèse n°2 d'une perception des variations de la 'surface visuelle' du figuré

La mise en doute de la première hypothèse nous conduit à soulever un problème quant à la signification de la quantité de noir représentée qu'il convient de résoudre. Elle conduit à examiner une seconde hypothèse selon laquelle l'œil percevrait d'abord une variation de la quantité de teinte, qui définit une surface visuelle.

Examiner cette seconde hypothèse n'est pas une chose aisée : elle nécessite de rendre proportionnelle la *surface visuelle*, donc la quantité de teinte, à la valeur du flux disponible dans la matrice afin de supprimer ce décalage, c'est-à-dire de donner une *largeur* de figuré égale à (F_{ij}/L_{ij}) , à une variable de rapport. Cette opération est illustrée par la Figure 6-5 dont on verra qu'elle conduit à une matrice d'intensité de flux qu'on note (I_{ij}) .

Figure 6-5. La construction d'une matrice d'intensité (I_{ij}) de flux



Cette hypothèse est en réalité intenable en cartographie de flux pour les deux raisons suivantes. La première raison est qu'elle conduirait soit à modifier de manière arbitraire la *longueur* du figuré – afin d'agir sur le produit *largeur*longueur* qui définit une *surface* – alors que cette dimension correspond justement à la traduction cartographique du concept de flux. Nous avons vu dans le chapitre 4 consacré à la sémiologie cartographique du flux, que le tracé du figuré illustre le concept même de flux, à savoir le *franchissement d'une limite de zone*, par la *longueur* du trait. Cette *longueur* est d'ailleurs disponible et donnée par le type de projection avant la *largeur*. Il n'est donc pas envisageable de la modifier, autrement que par le biais d'un changement de projection qui, nous l'avons vu, ne peut être une solution puisqu'il conduit à modifier l'information. La seconde solution consiste alors à agir sur la *largeur* du figuré, en la pondérant de manière à ajuster la *surface visuelle* à la valeur du flux, mais elle renvoie un résultat non exploitable à ce stade de la démonstration.

En effet, la mise en œuvre de cette seconde proposition conduit à générer une nouvelle valeur (Iij) issue du rapport entre la *largeur* et la *longueur* du figuré, afin que puissent coïncider la surface visuelle et la valeur du flux. La Figure 6-5 détaille la construction de cette nouvelle valeur. La *largeur* correspondant à la valeur (Fij) du flux, nous avons considéré la *longueur graphique* du figuré, exprimée en millimètres, en prenant l'exemple de la projection polaire pour construire (Iij).

Etant donné le type de résultats obtenus, à savoir une variable de rapport et non un stock, *l'information disponible* dans cette nouvelle matrice (les cases de la matrice qui définissent la *largeur* du flux) est impossible à représenter dans le cadre d'une *logique de flux*. C'est pourquoi l'hypothèse n°2 est, par conséquent et par essence intenable, pour des raisons d'ordre méthodologique et thématiques.

Sur le plan méthodologique, il n'est pas possible d'additionner ces différentes mesures (Lij/Dij) en ligne ou en colonne, tout comme il n'est pas envisageable de mobiliser directement la relation de proportionnalité de la *Taille* pour la cartographier (en agissant sur la seule *largeur* du figuré). Cette matrice d'intensité pose alors des problèmes de cartographie assez compliqués qui ne sauraient être résolus en ne mobilisant que la *Taille*. Nous verrons dans le chapitre 7 (section 7.1.1, page 288) que cette matrice (Iij) mobilisée avec la *Valeur* ou la *Transparence* pourra apporter une solution intéressante.

Sur le plan thématique, cette matrice – qui est analogue à celle d'un débit – possède en fait peu d'applications dans notre cadre d'analyse, car elle nécessite la prise en compte du temps (Tij). Mobiliser le temps signifie que l'on dispose de données de flux collectées à plusieurs dates, c'est-à-dire d'une matrice de flux découpée en périodes, ce qui n'est pas notre cas. Dans le cadre de cette thèse, nous avons fait le choix de ne mobiliser qu'une seule matrice de flux et non une série temporelle.

Nous en concluons que l'hypothèse n°2, pourtant **intéressante sur le plan théorique, conduit en pratique à une impasse liée à l'impossibilité d'agir sur la surface visuelle d'un objet qui est inexistant**, à ce stade de la démonstration.

Ce résultat nous conduit à éluder cette hypothèse n°2 et nous ramène à l'hypothèse n°1, dans l'objectif d'examiner dans quelle mesure il serait possible de mobiliser la *longueur cartographique* du figuré. Pour cela, il convient d'introduire *l'espace* qui, d'après l'hypothèse n°1, ne joue aucun rôle dans le processus de cartographie de flux. Est-ce vraiment le cas ?

La carte de flux étant réalisée dans un espace qui est toujours projeté, la *longueur du figuré* correspond en réalité à une *longueur cartographique* (Lij) – donc lue sur une carte projetée – qui, lorsqu'elle est perçue conjointement à la valeur (Fij), caractérise une *surface (visuelle) cartographique*, c'est-à-dire qu'il est possible de mesurer sur une carte, à l'aide d'une métrique.

Nous émettons alors une troisième hypothèse selon laquelle **l'œil percevrait une variation de surface du figuré qui serait liée, non pas à une quantité de teinte, mais plutôt au produit de la**

largeur par la longueur cartographique du figuré, à une surface mesurée dans une unité. Cette hypothèse n°3 nous amène à introduire la *longueur* en tant que dimension intrinsèque du figuré et pour cela, il est nécessaire de modifier notre raisonnement. Nous nous inscrivons alors dans le cadre de la *logique de mouvement* suggérée sur la Figure 6-1 (page 249).

En mobilisons ainsi l'espace, par le truchement de la *longueur cartographique*, nous souhaitons montrer qu'elle correspond peu ou prou à une acception de la *distance parcourue par le flux*. Pour cela, nous sortons du cadre strict de la carte de flux classique dont nous avons vu qu'elle faisait fi de la position des lieux. Pour distinguer le raisonnement que nous allons adopter par la suite, de celui de la *logique de flux* classique, nous introduisons la notion de **logique de mouvement** pour l'hypothèse n°3 que nous examinerons dans la section 6.3, située page 272. Avant cela, il est nécessaire de présenter les différentes *distances (parcourues par le flux)* qui sont mobilisables en cartographie de flux, car elles ne sont pas sans effet sur la perception et la signification de l'information représentée.

6.2 Les distances (Dij) mobilisables en cartographie de flux (Fij)

La troisième hypothèse de perception d'une variation de la surface du figuré nous conduit à introduire une donnée exogène : la *longueur* du figuré qui correspond à la *distance parcourue par le flux*, comme forme d'expression simplifiée de séparation spatiale, pour un système de projection donné. Les aspects purement méthodologiques, liés à la prise en compte de cette distance dans la cartographie des flux étant examinés dans la première section du chapitre 7, nous allons mobiliser ici cette notion de *distance parcourue* d'un seul point de vue cognitif. Le raisonnement proposé s'inscrit dans le cadre théorique défini par W. R. Tobler, conduisant à « [...] *une analyse symétrique des lieux et des flux à travers les concepts de mouvement et d'accessibilité* » (Grasland, 2009 : 5). En fait, l'analyse proposée s'appuie sur l'analyse de la perception de la distance représentée sur une carte de flux.

Comme précédemment, nous allons examiner la perception par un observateur quelconque, statique, de la distance sur le tracé du figuré.

La notion de distance est, en effet, très importante car sa prise en compte conduit à inscrire notre analyse dans un espace, plus précisément dans un territoire façonné par des sociétés. Trois familles de distances sont impliquées dans le processus de cartographie d'une matrice de flux :

- la *distance cartographique* (Lij) appréhendée comme la forme perçue sur la carte de la *longueur* du lien ;
- la *distance géométrique* (Dij) : elle est *géographique* et égale au kilométrage parcouru dans l'espace géographique, sur un réseau ; elle est souvent considérée sous une forme généralisée à la distance à vol d'oiseau (dite *euclidienne*), dans un besoin de simplification de la mesure géographique car l'espace géographique n'est pas isotrope ;
- la *distance empirique* ou *fonctionnelle* (Eij) correspondant à la distance temps, au coût généralisé de déplacement défini au regard de la thématique concernée par le flux.

Examinons ces différentes distances dans les sections suivantes.

6.2.1 La distance cartographique (Lij)

La première perspective d'intégration de la *distance parcourue par le flux* est d'ordre purement visuel. Elle est représentée par la *longueur* du lien, du trait ou de la bande qui illustre le flux. Elle est mesurée en millimètres sur le papier et correspond à une métrique réelle.

Nous introduisons cette notion de *distance cartographique* en écho à celle de « *distance graphique* » mobilisée par W. Tobler (1969), J.-C. Müller (1979) et A. L'Hostis (1996, 2003, 2008) dans les travaux déjà présentés dans la section 4.2.1 (page 179). Cette notion, mise en œuvre pour des graphes planaires, ne semble pas avoir fait l'objet, à notre connaissance, d'une application à la représentation des graphes non planaires spatialisés, qui correspondent à nos matrices de flux. Et pourtant,

l'utilisation conjointe d'une matrice de flux et d'une matrice de distance est susceptible de renouveler la figure de la carte de flux réalisée en tenant compte de la *distance parcourue par le flux*.

Sur une carte de flux, les dimensions du figuré sont toujours appréhendées par le cartographe *a priori* dans le cadre de la géométrie euclidienne : la *largeur* du figuré correspond à la valeur du flux – en application de la variable visuelle *Taille* – ; sa *longueur* est assimilée à la distance euclidienne qui sépare les lieux d'origine et de destination. Elle correspond à ce lien que nous représentons sous la forme d'un trait droit pour illustrer le caractère direct du transfert dans l'espace. Certains auteurs le représentent sous une autre forme, légèrement courbe, par exemple, en justifiant cette action par des raisons esthétiques. Si cela revient au même sur le plan de la construction du figuré, nous verrons (section 6.2.3, page 266) qu'il n'en est pas de même sur le plan de la perception.

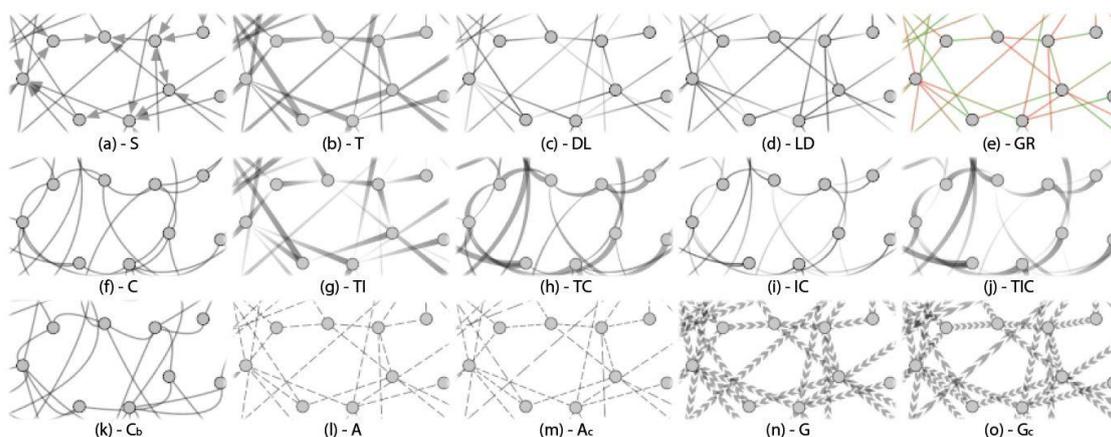
Cette posture est parfaitement justifiée à première vue : la *longueur* perçue du trait correspond bien à l'espacement des lieux et sa *largeur* à la valeur du flux (Fij). Cependant, l'examen de l'hypothèse n°2 nous a permis de démontrer l'impossibilité de faire concorder la surface du figuré de manière proportionnelle à la valeur du flux puisque cette éventualité est intenable : la position des lieux étant fixe dans un système de projection donné, il n'est pas envisageable de la modifier. Ainsi, il n'est pas possible de respecter *stricto sensu* le principe d'application de la *Taille*.

Cette pratique, parfaitement admise dans le registre de la cartographie de flux, est forte de conséquences. Elle signifie, en effet, que la méthode de cartographie classique de flux ne tient pas compte de l'existence de distances empiriques, c'est-à-dire de contraintes physiques et temporelles inhérentes au déplacement d'une quantité entre deux lieux. Si l'observateur perçoit bien un trait symbolisant une relation entre deux lieux, qui varie en *long* et en *large*, en fonction de la distance géographique et de la valeur du flux, celui-ci ne correspond pas exactement au choix du cartographe, au phénomène qu'il a souhaité représenter. La cartographie statistique ne fait pas partie, à notre connaissance, des disciplines qui postulent la fin de la distance (O'Brien, 1992) ou la platitude de l'espace (Friedman, 2005). La difficulté est donc liée au caractère absolu de la composante spatiale et à la nécessité de respecter un rapport de similitude convenable entre la sphère et le plan.

Les solutions apportées à ce problème de représentation de la distance empirique sur une carte de flux, l'ont été de deux manières. La première solution est d'ordre graphique, elle consiste à appliquer une transformation au tracé du trait pour l'incurver. La seconde est d'ordre mathématique : elle consiste soit à simplifier la figure, soit à la complexifier.

Les solutions graphiques sont appliquées soit à main levée, soit de façon (semi-)automatique, comme illustré sur la Figure 6-6.

Figure 6-6. Exemples de solutions liées à la transformation de l'apparence des liens

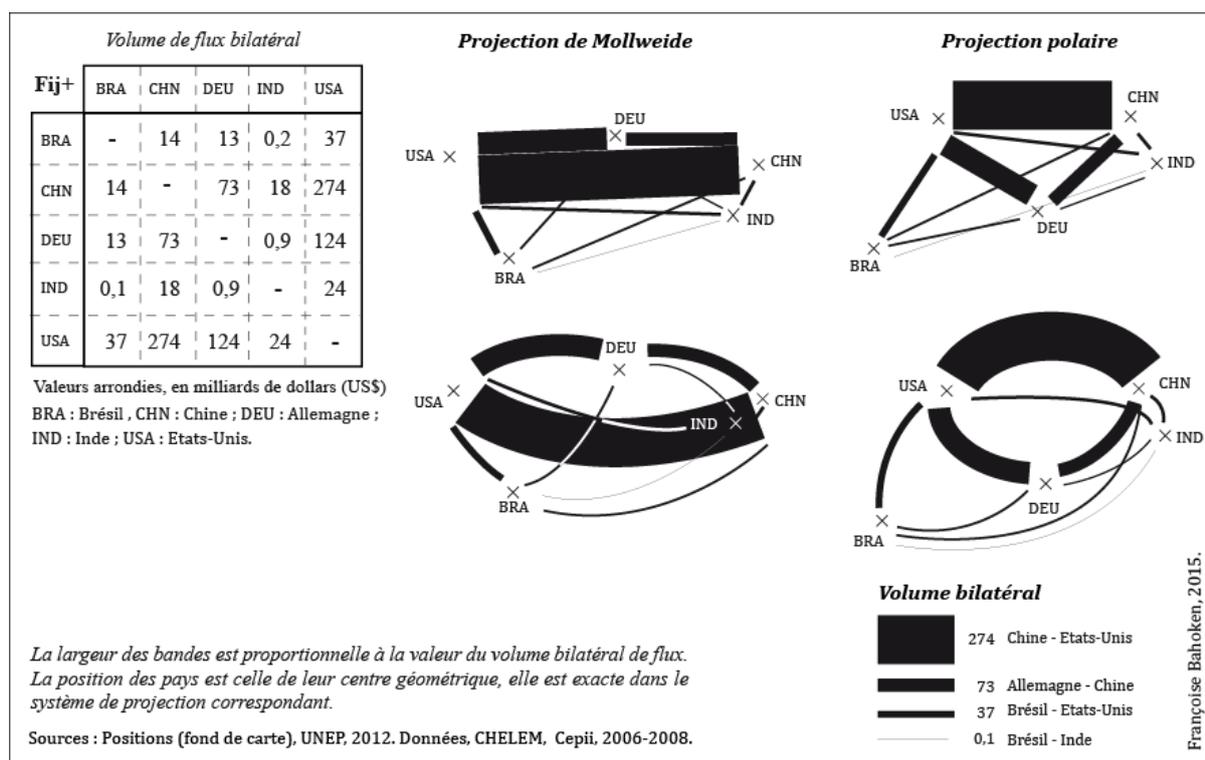


Source : D. Holten et al. 2010, (extrait du poster).

Ces modifications esthétiques de l'apparence du lien réalisées dans l'objectif de mieux représenter le *mouvement spatial* entraînent une modification de sa *longueur*. Par exemple, dans le cas de *l'edge bundling* (déjà mentionné dans la section 4.1.6, page 169), l'automatisation est censée faciliter la perception du mouvement à l'échelle locale. Elle a des effets sur la perception de la distance parcourue par le flux : le trait courbe apparaissant beaucoup plus long que dans la réalité, c'est comme si le flux ainsi perçu avait parcouru une distance plus importante, ce qui bien entendu n'est pas toujours le cas.

La Figure 6-7 met en évidence l'effet du changement de l'apparence (droite ou courbe) des liens sur la perception de flux internationaux.

Figure 6-7. L'effet du changement de l'apparence des liens et de la projection cartographique sur la perception des flux internationaux de marchandises



On voit bien sur les deux cartes de la Figure 6-7, réalisées avec un même système de projection, comment le changement de projection cartographique modifie la position des lieux et la *longueur cartographique* du trait qui relie les couples de lieux. On comprend également que cela n'est pas sans conséquence sur la perception de la distance parcourue par le flux. Lorsque le trait est droit, l'adoption d'une projection polaire entraîne une réduction apparente de l'information représentée que la courbure du trait ne rattrape que partiellement. La distance cartographique apparaît aussi plus courte, sauf lorsque le trait est courbe. La *longueur cartographique* du couple USA-CHN est de 53,18 mm, sur la projection de Mollweide, lorsqu'elle est représentée avec un trait droit et de 55,26 mm, sous la forme d'un trait courbe. De la même façon, sur la projection polaire, la *longueur cartographique* du couple USA-CHN est de 33,046 mm, sous la forme d'un trait droit et de 35,56 mm, lorsqu'est représentée avec un trait curviligne.

Cette variation de la longueur du trait n'est pas liée à notre exemple mais plutôt au fait que les flux considérés s'expriment à l'échelle mondiale, ce qui nous conduit à préciser ci-dessous (voir section 6.2.1.1 ci-après) *l'effet d'alignement topologique (ou de position des lieux)*, déjà évoqué dans la section 4.1.4 (page 158).

Le problème de l'utilisation de traits courbes tient ainsi au fait que les cartes de flux correspondantes traduisent une représentation bien plus erronée de la réalité que lorsqu'elles sont réalisées avec un trait droit, quand bien même la carte correspondrait à une représentation de cette réalité parmi d'autres.

On observe ainsi que la cartographie de flux réalisée à l'échelle mondiale pose en réalité un double problème : le premier est un *effet topologique de position* des lieux (ou de *projection*, lorsqu'il est appréhendé à l'échelle mondiale) et qui impose une réflexion sur le choix d'un système de projection cartographique adapté au processus que l'on souhaite montrer ; le second problème est théorique et conceptuel, car il contraint, dans certains cas lié au type de flux, la représentation des routes empruntées par le flux (voir section 6.2.1.2). C'est *l'effet d'itinéraire*. Ces deux effets sont respectivement définis dans les sous-sections suivantes.

6.2.1.1 *L'effet d'alignement topologique, un problème spécifique de cartographie de flux à l'échelle mondiale*

La cartographie des matrices de flux qui s'expriment à l'échelle mondiale est caractérisée par un problème particulier lié à la situation des lieux, nous allons l'évoquer ici mais sans le développer. En fonction de la projection cartographique utilisée, la distribution spatiale des lieux se caractérise, en effet, par ce que l'on qualifie d'*effet d'alignement topologique* des lieux qui justifie un arbitrage dans le choix du système de projection cartographique à utiliser, au regard de la thématique.

Le problème résulte du fait que la cartographie de flux mondiaux conduit à la représentation d'un phénomène lié au processus de mondialisation mettant en évidence la Triade, ou désormais les BRICS. La mondialisation ne concernant en réalité que quelques lieux dits centraux, qui sont majoritairement situés dans l'hémisphère nord, les cartes qui en résultent sont souvent biaisées par un *effet d'alignement horizontal*, qui est préjudiciable à la perception des centralités et de la hiérarchie des lieux. Les solutions qui sont apportées pour résoudre ce problème appartiennent à la famille des transformations cartographiques de position, parmi lesquelles les changements de système projection. C'est la raison pour laquelle les fonds de carte traditionnels de type Mercator ou Mollweide s'avèrent en fait inadaptés à la représentation de certains processus mondiaux, du fait de cet effet graphique, et aussi en raison de la vision du monde qu'ils entraînent.

La question du choix de la projection à utiliser à l'échelle mondiale est très ancienne. Elle a fait l'objet de nombreuses discussions qui se sont toutes inscrites dans la recherche d'un optimum supposé de représentation de la Terre. Elle a aussi donné l'opportunité d'examiner la manière dont différents auteurs ont interprété le centre du monde ou bien le pouvoir sur le monde, en le plaçant au centre de la planche. C'est d'ailleurs l'une des raisons pour lesquelles la réflexion sur la meilleure manière de représenter des flux à l'échelle globale puis mondiale, se développe dans le contexte de la première Guerre mondiale (voir chapitre 4). C. Grataloup et F. Durand-Dastès (1992) ont d'ailleurs bien montré le rôle des flux internationaux dans la structuration de l'espace mondial, selon le modèle de la Triade. « *C'est à cette période [dans les années 1980] que les types de planisphères se sont multipliés, traduction de cette prise de conscience de l'obsolescence, au moins partielle, des visions cartographiques du monde [notamment de la centralité européenne]* » (Grataloup, 2011 : 5). Cependant, s'ils permettent de diversifier la vision du monde (Grataloup, Fumey, 2015), la persistance du débat quant à la meilleure projection cartographique à utiliser est vaine. Cela conduit d'ailleurs C. Grataloup à conclure à « *l'impossible planisphère* », à l'issue d'une expérimentation menée sur une matrice de flux internationaux de marchandises par voie maritime. « *On vérifie ainsi qu'il n'y a pas de planisphère parfait. Pourtant, chacun présente des avantages. Il importe surtout de multiplier les types de planisphères, quitte à perdre un peu de temps à chaque nouveauté, mais pour envisager la terre sous le plus d'angles possibles* » (Grataloup, 2011 : 62).

6.2.1.2 L'effet d'itinéraire

L'effet d'itinéraire est le second problème de la cartographie de flux réalisée à l'échelle mondiale, il correspond à la nécessité de représenter les routes empruntées par le flux au risque d'interprétations erronées.

Dans l'analyse de différents systèmes de projection cartographique qu'il propose, C. Grataloup examine plusieurs visions des flux mondiaux de marchandises, en suggérant plusieurs alternatives à la projection de G. Mercator (1569) : une projection de Bertin, une projection polaire et la projection Atlantis, sous la forme de « *Boîte à outils* » (Grataloup, 2011 : 58-59 ; 62-63). Cela conduit l'auteur à mentionner un second problème, déjà identifié lors d'une précédente analyse réalisée avec F. Durant-Dastès et spécifique à la thématique du flux représenté à l'échelle mondiale : la représentation des routes empruntées.

La seconde particularité de la cartographie de flux mondiaux tient dans la nécessité de décrire les itinéraires ou routes empruntées, en l'occurrence des flux de conteneurs réalisés par voie maritime qui, pour des raisons pratiques et liées à l'échelle mondiale de représentation, apparaissent toujours sur la carte comme étant transportés par voie terrestre. En l'absence d'informations sur la trajectoire exacte des navires, les flux représentés par des liens droits franchissent toujours peu ou prou des masses continentales, alors qu'ils devraient rester cantonnés sur les océans ; il ne saurait d'ailleurs en être autrement. C'est pourquoi C. Grataloup conclut dans son expérimentation cartographique, sur le caractère erroné des trois cartes mondiales des flux commerciaux présentées, aucun des fonds de cartes ne paraissant adapté.

Outre la correction éventuelle de ces deux effets, la seconde famille de solutions apportées au problème de la perception de la distance parcourue par le flux qui minimise les biais, est d'ordre mathématique. Les solutions correspondantes ont été examinées exclusivement dans le cadre de la théorie des graphes par W. Tobler (1969), J.-C. Müller (1979) et A. L'Hostis (1996, 2003, 2008), à travers la notion de « *distance graphique* », qui correspond à la *longueur* du lien perçue, à notre *distance cartographique*. Il nous semble important de revenir sur ces notions que nous avons déjà évoquées dans le chapitre 4, afin de situer notre position par rapport aux travaux de ces auteurs.

Nous avons montré (voir section 4.2.1.2, page 180) comment la distance graphique pouvait être transformée de manière à représenter une distance-temps entre deux lieux, en fonction de différents modes de transport, comme dans le cas des « *chronocartes* » (L'Hostis, 1996) dont il a déjà été question. Cette proposition a ouvert la possibilité d'utiliser la troisième dimension de l'espace pour figurer la distance parcourue, là où W. Tobler (1969) et J.-C. Müller (1979) avaient focalisé leurs efforts sur la deuxième dimension, en modifiant la *longueur* du tracé : c'est le formalisme des « *cartes en ressort* ». Il est intéressant de noter que les solutions proposées ne sont pas dans le registre de l'anamorphose vectorielle, car elles n'affectent pas la position des lieux, mais seulement la *longueur* ou la *hauteur* du lien.

Les travaux de W. Tobler et de J.-C. Müller participent en particulier à la recherche de solutions au problème de la représentation de la distance parcourue par un flux, avec les deux contraintes suivantes :

- l'utilisation de la *longueur* du figuré, correspondant à une acception de la distance géographique ;
- le recours à un tracé linéaire correspondant au référentiel euclidien, en supposant que les distances cartographiques et géographiques concordent.

D'après J.-C. Müller (1979), le problème de la représentation linéaire de la *distance parcourue* tient à cette double contrainte de représentation de la distance par un trait droit, dans un référentiel euclidien qui correspond au trajet direct réalisé à vol d'oiseau.

Le fait est que le tracé des flux présents sur la carte ne correspond « jamais » à la distance réelle parcourue. Il ne coïncide pas non plus avec la morphologie du déplacement qui emprunte toujours un réseau de transport, soit un trajet sinueux. Le cheminement parcouru par le flux est de fait soumis à la rugosité de l'espace, ce qui le rend par essence non linéaire. C'est pourquoi sa représentation linéaire, sous la forme d'un trait droit, peut ainsi être considérée comme abusive, car elle ne correspond pas à la réalité.

La perception de la distance représentée sur une carte coïncide avec la notion de « *distance graphique* » (Müller, 1979 : 224), supposée traduire sur la carte la *distance euclidienne* et cela, de manière équivalente (ou conforme) à la *distance géographique* (kilométrique), moyennant le rapport d'échelle (ce qui est généralement le cas). Le cartographe a souhaité matérialiser cette distance. Cependant, sa perception correspond à une construction mentale, puisqu'elle est établie à partir de la distance euclidienne qui ne correspond pas toujours à la réalité : par exemple, les migrations de populations humaines ne s'exprimant pas (sauf erreur), en ligne droite comme celles des oiseaux migrateurs.

Il en résulte que la *distance cartographique*, telle qu'elle est mise en œuvre en cartographie de flux, renvoie donc à **une information doublement généralisée** : elle est liée au choix de la distance euclidienne (qui ne correspond pas à la réalité) et à la transcription graphique d'un déplacement sinueux sous la forme d'un trait droit.

La mobilisation de la *distance cartographique*, dans un référentiel euclidien, pour cartographier des flux, n'est toutefois pas un problème en soi puisque dans une matrice de flux, le cheminement exact entre les lieux n'est pas connu avec précision. La mesure portant sur la valeur d'un lien direct entre des OD, il n'est pas absurde de le représenter sous la forme d'un trait droit, l'espace n'étant pas pris en compte dans la mesure. Pour qu'il le soit, il faudrait que la mesure du flux prenne en compte la distance parcourue ; par exemple, dans le champ du transport où il est question de mouvements de marchandises, ces derniers sont exprimés, parfois, en unités de tonnes.kilomètres (voir chapitre 1).

De plus, nous verrons dans la section suivante que la métrique euclidienne à laquelle cette longueur renvoie, bien que n'étant pas des plus précises, est considérée par différents auteurs parmi lesquels W. Tobler (1969) comme une bonne approximation de la *distance géographique* (Dij). En conséquence, si le fait que le trait qui relie un couple de lieux soit droit ou non ne change pas fondamentalement l'information représentée (Fij), le problème reste qu'il modifie sa perception, ce qui n'est pas anodin. Ainsi, un changement de l'apparence de la figure, par l'incurvation d'un tracé, comme dans l'exemple Figure 6-7 (page 260), entraîne une modification de l'information représentée, de la même façon qu'un changement de projection modifie la perception du processus cartographié. Cela tient au fait que l'introduction d'une courbure du lien entraîne une augmentation de la *surface* du figuré, par conséquent de la distance entre les lieux puisque la *largeur* du figuré reste fixe. C'est pourquoi dans l'hypothèse n°3 où l'œil percevrait la *surface*, l'information reçue est erronée.

Cette proposition ne sera pas davantage examinée dans le détail dans le cadre de cette thèse, car elle fait partie d'une famille de solutions liées aux transformations (carto)graphiques, qui entraînent une modification de la perception du message.

La perception de la *distance cartographique* variant avec la configuration spatiale, nous allons supposer qu'elle correspond à la *distance géométrique* mesurée entre les barycentres des mailles.

6.2.2 La distance géométrique : géographique ou euclidienne (Dij)

La *distance géométrique* est la mesure de l'espacement des lieux d'une zone d'étude. Elle correspond en réalité à une approximation de la *distance géographique* qu'il est nécessaire de parcourir pour rejoindre deux lieux d'OD. Cependant, le fait est que dans un monde abstrait (ou dans un espace physique théorique), la figuration de la distance à parcourir entre deux lieux est généralement

abstraite. Elle est mesurée sur un plan ou sur une sphère à vol d'oiseau. Lorsqu'elle est transposée à la surface de la terre, on observe que cette métrique fait fi de l'occupation terrestre ou maritime de l'espace d'étude, susceptibles d'affecter sa mesure.

La Figure 6-8 illustre différentes manières de calculer la distance géométrique entre deux points, selon qu'elle est effectuée sur un réseau routier (comme dans le cas d'une distance géographique), par la mer en longeant la côte (voie maritime) ou par avion, à vol d'oiseau.

Figure 6-8. Illustration des distances géométriques : géographiques et euclidiennes (Googlemaps, 2015)



À l'échelle locale, la distance géométrique est figurée par un trait droit : elle correspond à une *distance euclidienne* mesurée à vol d'oiseau, considérant le *plus court chemin géométrique*. Le *plus court chemin géographique* s'en écarte, car il est mesuré sur un réseau de transport et exprimé dans une unité de distance géographique : en kilomètres – comme la mesure euclidienne – mais aussi en temps de transport. Sur la Figure 6-8, la distance à parcourir entre Le Havre et Les Pieux est soit de 150 km environ à vol d'oiseau contre 223 kilomètres par la route, ce qui correspond à un temps de parcours de 2h21 dans un véhicule individuel ; sachant que cette distance aurait aussi pu être mesurée par voie maritime ou aérienne, à l'aide d'un drone, par exemple (l'aéroport le plus proche des Pieux étant celui de Caen-Carpiquet).

La mesure géographique d'un déplacement varie en fonction du milieu. Elle est soit terrestre, soit maritime, soit aérienne, ce qui signifie qu'elle est susceptible de varier en fonction de contraintes géographiques liées au type d'espace traversé (terre, mer, air), mais aussi en raison de contraintes topographiques et liées aux conditions de transport. Par exemple, dans le cas d'un déplacement par voie terrestre, celui-ci longe la côte et utilise des tronçons de voies routières qui sont différentes dans leurs caractéristiques (Nombre de voies, vocation, nombre de chaussées) et dans l'espace traversé (terrain plat ou légèrement pentu, sinueux ...). Ces différences introduisent une contrainte sur la vitesse de circulation qui affecte *in fine* la mesure temporelle de la *distance géographique* et non sa mesure kilométrique qui reste fixe. Sur la Figure 6-8, pour rejoindre Les Pieux depuis le Havre, on circule, en effet, plus rapidement sur la portion de voie rapide située entre Le Havre et Caen que sur la route nationale n°12 nichée dans le bocage normand (!).

La *distance euclidienne* correspond ainsi à un cas théorique d'utilisation de la *distance géographique* parcourue par le flux. Elle entraîne une approximation à vol d'oiseau de la mesure exacte d'un cheminement réalisé sur un réseau. C'est pourquoi elle est toujours exprimée dans une unité de mesure, le kilomètre, par exemple. Sur le plan graphique, alors que la *distance géographique* correspond à une représentation sinueuse, elle est fréquemment approximée par la *distance euclidienne* qui correspond au trait droit. Cela ne pose pas de problème à l'échelle locale, comme nous verrons sur

la Figure 6-10 (page 271), dans la mesure où la distance géographique correspond, sur le plan théorique, à une mesure du plus court chemin entre deux points situés à la surface de la terre, généralement exprimée en kilomètres.

La *distance géométrique* peut ainsi être interprétée de deux manières, en fonction de l'échelle : comme une *distance géographique*¹⁸⁰ ou comme une *distance euclidienne*.

Plus généralement, intégrer une mesure d'éloignement géographique correspond au cas où la matrice des données de flux est associée à une matrice des positions géographiques $P(x,y)$, caractérisant le système de coordonnées des lieux d'origine et de destination. Considérer ce système de points revient en pratique à faire le choix d'un système de projection dans lequel sera représentée la carte. La *distance géographique*, entre les points projetés sur le plan, est généralement formulée à partir de couples de coordonnées (x,y) qui sont supposés représentatifs des lieux d'origine et de destination (il s'agit souvent des barycentres). La mesure obtenue est nécessairement double : elle fait référence, d'une part, à la *distance géographique* (D_{ij}) qui correspond à la mesure réalisée sur la sphère – à petite ou à grande échelle – et d'autre part, à une mesure de *distance cartographique* réalisée sur le plan et qui sera homologuée à la distance à vol d'oiseau dite *euclidienne*.

La *distance euclidienne* est souvent calculée faute de mieux à partir des couples de coordonnées (x,y) des lieux d'origine (i) et de destination (j) : ils correspondent soit à un lieu particulier, soit à leur barycentre (ou centroïde).

Si les mesures de *distances géographiques* et *cartographiques* sont globalement similaires dans les cartes réalisées à grande échelle, l'altération des mesures est d'autant plus importante que l'échelle devient petite. Cela explique le fait que la *distance géographique* soit assimilée à la *distance euclidienne* dans les analyses locales. Aussi l'importance du choix du système de projection pour les cartes réalisées à l'échelle mondiale, par rapport à celles qui sont régionales ou locales.

Proposer une cartographie de flux à grande échelle, dans un espace projeté qui altère le moins possible les distances à vol d'oiseau, permet d'introduire implicitement des hypothèses sur le rôle de la distance dans la structuration de l'espace géographique. Le fait que la *distance cartographique* corresponde à la *distance géographique* – avec $D_{ij}=k(L_{ij})$ compte tenu du rapport d'échelle $(1/k)$ – si et seulement si la projection est équivalente et sans croiser le bord de la carte – sous-entend une position théorique en faveur de la première loi de la géographie de W. Tobler (1970) : « *J'évoque la première Loi de la géographie : toutes les choses sont reliées entre elles, mais celles qui sont proches le sont davantage que celles qui sont éloignées* » et, plus généralement, des modèles d'interaction spatiale de type gravitaire. En effet, de ce point de vue, une carte de flux dont les positions des lieux sont déduites de celles des flux conduit, pour un système de projection donné, à éloigner les lieux selon une certaine mesure de distance comme l'a montré (Grasland, 2009 : 6-7) ; cette distance étant toujours supposée avoir un effet négatif assimilé à un frein au déplacement (Tobler, 1970 ; Poulain, 1981, 1991).

A contrario, **un graphe de flux** (voir chapitre 2, Figure 2-9, page 69) **où les positions des lieux sont sans aucun rapport avec leurs positions géographiques suppose une hypothèse théorique forte de fin de la distance ou de mort de la géographie**. La position de ces géographes et sociologues (O'Brien, 1992 ; Friedman, 2005), considérant l'espace sans relief, devient dès lors à la fois logique et compréhensible. Si ces derniers devaient représenter graphiquement les interactions qu'ils analysent, ils le feraient très probablement sous la forme de matrice-ordonnée ou de graphe, c'est-à-dire sans

¹⁸⁰ A grande échelle ou pour l'analyse des mobilités régionales ou locales, la *distance géographique* est parfois assimilée à la distance réseau, c'est-à-dire au trajet réalisé sur une infrastructure de transport réseau particulier ou sur un maillage régulier, comme dans le cas des villes américaines. Elle correspond donc à un trajet réel effectué dans l'espace-temps, ce qui rejoint la définition de la distance empirique que nous allons examiner dans la section 6.2.3.

tenir compte de l'espace, plutôt que de cartes de flux. Cela étant, dans un pur espace de flux, la distance entre les lieux n'en sera pas pour autant arbitraire. Elle impliquera nécessairement d'autres lectures des proximités qui peuvent être sociales, économiques, culturelles, ...

L'intérêt du recours à la *distance géographique* dans l'analyse et la cartographie des interactions spatiales est par conséquent indéniable. Ses propriétés de symétrie et d'inégalité triangulaire rendent par ailleurs son usage aisé tant d'un point de vue numérique que cartographique. Elle correspond aussi à un bon substitut de la distance réelle, empirique.

6.2.3 La distance empirique (Eij)

Dans la typologie¹⁸¹ des distances qu'il propose, J.-C. Müller (1982) qualifie la *distance empirique* de « *distance fonctionnelle* » aussi nommée « *distance d'effort* » dans le champ des transports. La distance empirique (Eij) correspond à la mesure d'un déplacement réalisé en tenant compte du mode de transport (par voie routière, ferroviaire, aérienne, ...). Elle est exprimée en unités kilo(métriques), temporelles (distance-temps) ou financières (en coût de transport). Lorsqu'elle porte sur des mobilités spatiales humaines, cette mesure est effective ou bien perçue par les individus enquêtés : le coût d'un déplacement pouvant, en effet, varier selon qu'il est estimé sous une forme monétaire ou perçu comme étant onéreux ou non, pénible ou agréable, ou encore en termes de possibilité de relation (de contact physique, d'information sur la destination...).

La distance empirique correspond aussi, dans une perspective économétrique, au coût généralisé de transport ou de déplacement dont la définition intègre fréquemment le temps. Elle est connue pour faire obstacle au flux, en le réduisant par des *effets de barrière* que les méthodes de modélisation des flux essaient d'estimer. L'introduction d'une distance empirique conduit à qualifier les obstacles aux relations qui sont spécifiques à la nature du flux considérée. La valeur de cette distance étant, en effet, susceptible de varier fortement, pour un même type de flux, en fonction de la topographie de l'espace qui contraint le trajet réalisé entre les lieux d'origine et de destination, ainsi que du mode de transport utilisé pour le déplacement (voir Encart 6-1).

Encart 6-1. Sur la variation de la mesure de la distance empirique (Eij)

« On 'sait' bien que 'la géographie' ne s'occupe que de faits bruts et évidents, qui ne donnent pas matière à penser. Pourquoi donc parler de la distance ? Pourtant, si vous mesurez sur une carte, en ligne droite et donc 'à vol d'oiseau', Tours n'est qu'à 194 km de Rennes. Prenez le trajet routier le plus direct par Sablé, ce sera 211 km, et il y faudra 3 h 30. Préférez l'autoroute et passez donc par Le Mans, la distance monte à 252 km, et c'est plus cher, mais vous mettez bien moins de temps (2 h 36 vous dit votre logiciel préféré). Et en train alors ? Plus personne ne vous indique les distances, et de Tours à Rennes la SNCF vous fait passer par Paris en 4 h 06 ou 5 h 18, à la rigueur par Nantes et Redon en 4 h 13, à des tarifs variant entre 40 euros et 82 euros en seconde classe, un jour normal sans billet de faveur pris six mois à l'avance, le plus cher n'étant pas le plus rapide. Encore n'évoquerai-je pas le canot de plaisance, l'avion ni la randonnée pédestre, équestre ou cycliste, qui fourniraient d'autres mesures d'arpentage, de temps et de coût ; ni l'ADSL, qui les réduit à (presque) rien. Alors quelle distance ? Des kilomètres, ou du temps, ou de l'argent ? Lequel économiser ? Et pour quoi faire ? »
Source : R. Brunet, 2009, pp. 14.

¹⁸¹ J.-C. Müller (1982) évoque les « *distances globales* » (qui correspondent aux distances physiques), les « *distances fonctionnelles* » et les « *distances cognitives* » ou « *métaphoriques* »

Si l'on considère, à titre d'exemple, les navettes de l'Isère, le recours à la distance-temps permet de considérer l'effort fourni par les enquêtés pour se rendre sur leur lieu de travail depuis leur lieu de résidence. Sur le plan thématique, cette nouvelle matrice de flux, construite en fonction des distances-temps de parcours, permet de mesurer l'éventuelle pénibilité de ces mobilités professionnelles ou encore la difficulté de se rendre à son lieu de travail, dans des conditions normales de circulation.

Si l'on considère les échanges commerciaux, l'introduction d'une matrice de prix de transport permet d'interpréter les flux de marchandises en termes de coût financier de transport, selon le type de biens concerné (c'est-à-dire selon leur valeur), de la route qu'ils ont emprunté et éventuellement, du coût environnemental de leur transport, au regard de leur bilan carbone ; la quantité de flux pouvant être exprimée en équivalent polluant de type CO₂.

La mesure d'éloignement considérée dans les deux exemples précédents est empirique ou thématique, elle correspond à la distance réelle parcourue par l'individu – la route empruntée par le travailleur qui se rend à son lieu de travail ; la voie maritime de la marchandise qui est acheminée vers son aire de marché – qui diffère de la distance traditionnelle à vol d'oiseau. De ce fait, elle ne devrait pas être représentée sous la forme d'un trait droit, comme la *distance euclidienne*, c'est-à-dire indépendamment des contraintes physiques, temporelles et monétaires qui la caractérisent. C'est parce que l'espace réel n'est ni homogène, ni isotrope que certains auteurs (Tobler, 1969, L'Hostis, 2003) ont souligné l'inadéquation du trait droit pour représenter la distance parcourue. L'utilisation d'un trait sinueux permettrait de corriger le biais visuel introduit par le trait droit mentionné par J.-C. Müller (1979), car il est associé à la métrique euclidienne du trajet le plus direct. D'après A. L'Hostis, « *il faut insister sur le fait que la distance ainsi définie est purement mathématique : elle est uniquement la mise en correspondance de couples de lieux avec une mesure qui prend la forme d'un écart. C'est un objet mathématique qui caractérise les données géographiques mais qui n'a pas a priori de traduction immédiate et univoque dans la sphère des représentations graphiques* » (L'Hostis, 2003 : 218). Les « *chronocartes* » de L. L'Hostis et autres « *cartes en ressort* » de W. Tobler représentant l'espace des flux de manière déformée, un apprentissage à la reconnaissance des formes proposées devient indispensable. Leur mise en œuvre n'est pas aisée, la perception de l'information ne l'est pas plus.

La cartographie de la *distance empirique* pose une réelle difficulté liée à la forte variabilité de sa mesure selon les lieux, les moyens de transport utilisés, les trajectoires utilisées mais aussi selon les individus. Le recours à une métrique empirique permet toutefois de relativiser la position respective des lieux les uns par rapport aux autres, compte tenu des perceptions de ces distances. C'est pourquoi les « *distances cognitives* » ou « *métaphoriques* », évoquées par J.-C. Müller (1982), sont fondées sur la perception de la proximité ou de l'éloignement du lieu de destination par un individu. Elles peuvent être appréhendées comme une mesure des contacts sociaux entre les lieux ou les individus, exprimée en termes de fréquence de déplacements selon un motif (professionnel, de loisirs,...), de nombre d'appels téléphoniques ou encore d'échanges culturels (Müller, 1982 : 190).

Faute de pouvoir donner une définition unique des mesures de *distance empirique*, nous pouvons cependant les classer selon qu'elles sont binaires, ordinales ou quantitatives continues. Nous y reviendrons dans le chapitre 7. La *distance empirique* étant par ailleurs généralement inconnue, il n'est *a priori* pas possible de la définir par des considérations statistiques ou cartographiques. Dans l'analyse des flux dont il est question, seul un thématicien serait en mesure de savoir si le temps importe plus ou moins que le coût de transport ou que l'énergie dépensée.

Les propositions de W. Tobler (1969) et A. L'Hostis (1997, 2003) pour cartographier la *distance empirique* conduisent à une déformation importante du tracé, certes proche de la réalité du terrain, mais difficilement lisible, qui a très probablement freiné la généralisation de leur usage. En effet, elles supposent un apprentissage visuel important. Face à ce constat, nous faisons le choix de conserver le

formalisme du trait droit et du rendu de la carte en deux dimensions. Par ailleurs, si la perception de la distance fonctionnelle est à peu près fixe à l'échelle locale, elle varie considérablement à l'échelle mondiale pour deux raisons. La première est liée à la sphéricité du globe qui contraint le tracé du plus court chemin « à vol d'oiseau » d'apparaître sous une forme légèrement courbe, suivant la rotondité de la terre. La seconde raison et non la moindre est liée à la thématique du flux.

A l'échelle mondiale, les flux transférés par voie maritime ou aérienne n'empruntent pas les mêmes routes, lesquelles ne correspondent pas, de surcroît, à des trajets directs. Elles ne devraient donc pas être cartographiées de la même manière, sous la forme d'un trait droit. Ces routes maritimes ou aériennes sont structurées par le passage obligé en quelques points stratégiques du globe, qu'il convient de représenter sur la carte.

L'Encart 6-2 précise les caractéristiques du cheminement de marchandises à l'échelle mondiale.

Encart 6-2. Sur le cheminement de marchandises à l'échelle internationale

En réalité, le cheminement d'un bien à l'échelle internationale varie en fonction de la valeur de la marchandise (Nierat, 2015)¹⁸². Cette assertion nous conduit à l'hypothèse provisoire selon laquelle la carte est susceptible de varier en fonction du type de flux (chapitre 1), selon qu'il s'agit de biens onéreux (des produits électroniques) ou non (de la ferraille) mais aussi de sa valeur monétaire. Par exemple, si l'on considère le transport maritime formé en partie de conteneurs de produits de valeurs et non onéreux (de ferraille) entre Hong-Kong et New-York, trois trajectoires sont envisageables. La première est classique : elle consiste à traverser l'océan pacifique pour atteindre les ports de l'ouest des Etats-Unis, puis de terminer le trajet en train pour rejoindre New-York. La seconde trajectoire est moins courante mais tout à fait possible : elle redescend l'Océan pacifique vers l'Indonésie, franchit le détroit de Malacca, le canal de Suez avant de traverser la méditerranée d'Est en Ouest pour rejoindre l'Atlantique, par le détroit de Gibraltar, puis le port de New-York. La troisième trajectoire est la plus rapide : depuis Hong-Kong, elle traverse l'océan pacifique en direction du canal de Panama, puis remonte en bateau jusqu'à New-York. Elle est cependant peu utilisée car les écluses de Panama ne permettent pas (encore) le passage des porte-conteneurs les plus importants en taille.

Ainsi, pour un même transport, trois trajectoires sont en concurrence pour ce qui est du temps du transport, du coût du transport et de la distance kilométrique parcourue. Le prix des marchandises transportées étant fonction de différents paramètres qui sont notamment liés aux déséquilibres du trafic, c'est-à-dire au taux de remplissage du conteneur. Si le bateau est plein de conteneurs, ces derniers ne le sont pas forcément, certains d'entre eux pouvant être vides et présenter une faible valeur par rapport à ceux qui sont pleins. Le transport de faibles valeurs suivra ainsi un itinéraire allongé par rapport à celui des fortes valeurs (leur cheminement plus court permettant de gagner du temps, donc de l'argent en réduisant l'immobilisation financière).

De ce point de vue, représenter ces flux de marchandises sous la forme d'un trait droit conduit au biais visuel déjà mentionné dans la section 6.2.1, page 258, au sujet de la *distance cartographique*. Pour l'éliminer, il convient de lui substituer un trait courbe ou sinueux – suivant le réseau ou les routes réellement empruntées – de manière à ce que les tracés ne franchissent pas des espaces interdits, par exemple, que les flux maritimes ne traversent pas de continents. En forçant ainsi le tracé à ne pas

¹⁸² Ces informations sont issues d'une conversation informelle avec l'auteur en 2015, au sujet du transport international de marchandises à l'échelle mondiale.

franchir d'espaces terrestres, cela revient à représenter les routes¹⁸³ réelles ou supposées empruntées par les mobiles, à représenter les flux sous une forme affectée sur réseau (maritime), à représenter un mouvement (de navires). **A l'échelle mondiale, la cartographie de flux maritimes contraint à l'application d'angles au tracé du figuré, lequel est dès lors représenté tel un cheminement : la carte de flux décrit alors un mouvement.**

La Figure 6-9, située page suivante, représente des flux internationaux de conteneurs réalisés par voie maritime, ils concernent la façade asiatique orientale et le reste du monde. Ils sont présentés selon deux systèmes de projection (Mollweide et polaire) et selon deux type de figurés : droits et sinueux.

La représentation des routes approximatives¹⁸⁴ empruntées par les navires (trait sinueux), sur les deux fonds de cartes, permet la figuration de leur mouvement. Ce mouvement correspond à la matérialisation du cheminement des navires entre les différentes façades maritimes, dans le monde réel. Il est intéressant de noter que la représentation des mouvements mondiaux correspond à un retour vers la proposition désormais historique de C.-J. Minard (voir section 4.1.2, page 146) qui n'est, hélas, plus en vogue, alors même que c'est celle qu'il convient sans doute d'adopter lorsque le cheminement des flux est connu. Puisqu'il en est ainsi, nous indiquons en bibliographie trois boîtes à outils pour ArcGis 10x permettant de générer des cartes distributives.

La Figure 6-9 met ainsi en évidence la variation de la perception des quantités en fonction du type de projection et de tracé du figuré et cela, en lien direct avec la thématique à savoir : la mondialisation des flux de conteneurs par voie maritime.

Cette variation du motif de ces flux mondiaux signifie qu'en réalité, à l'échelle mondiale, la perception des flux est susceptible de varier non seulement pour des raisons esthétiques ou cartographiques (comme dans le cas d'un changement de projection) mais surtout en fonction de la thématique dont il est question.

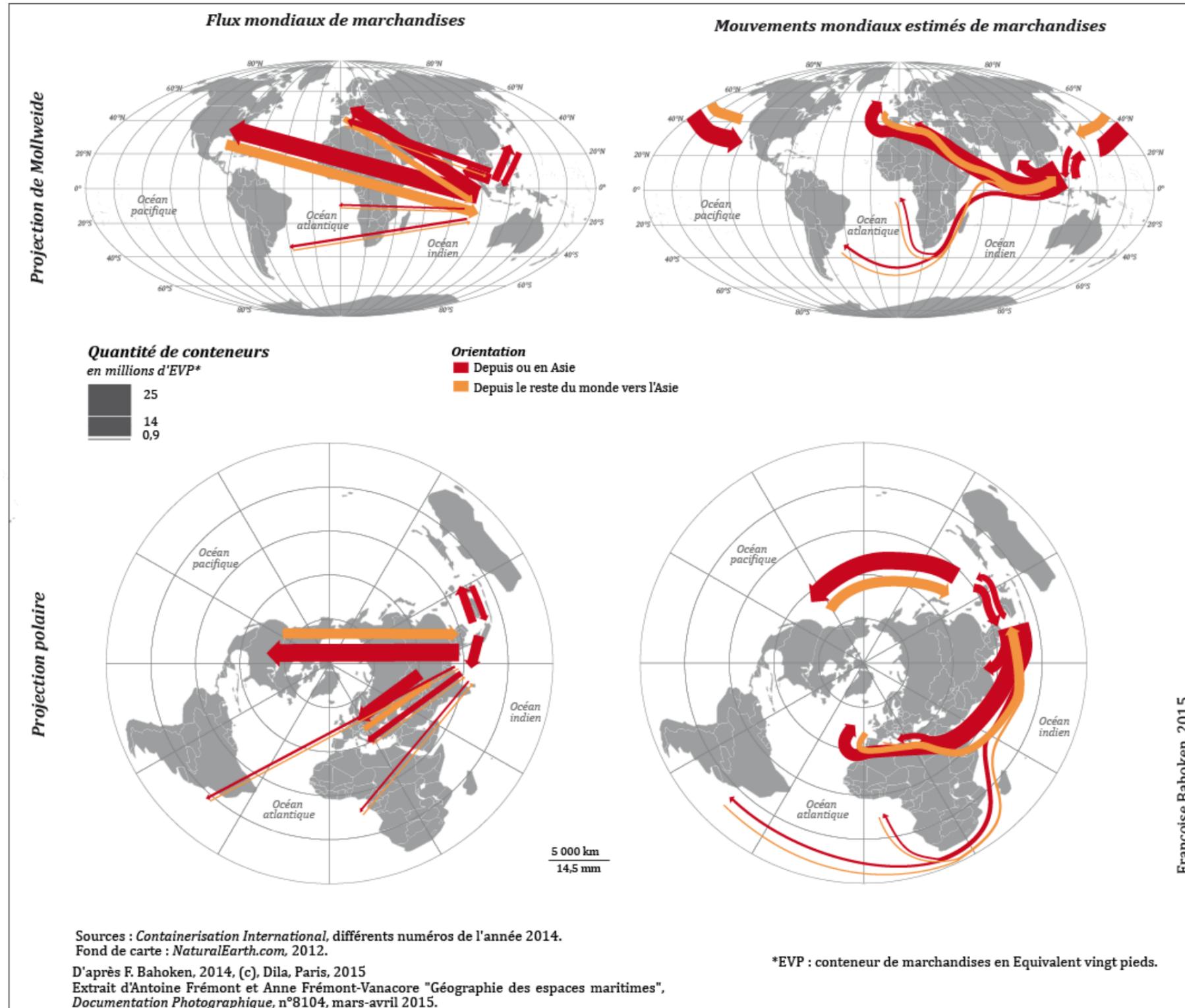
Cependant, la représentation de manière sinieuse, sous la forme de trajectoire, de données de flux Origine-Destination, par essence directs, peut apparaître contradictoire puisque le cheminement de ces derniers n'est pas connu avec précision. On est ainsi en droit de s'interroger sur la pertinence du recours à la cartographie de flux pour représenter un mouvement : est-ce l'objectif d'une carte de flux que de représenter un mouvement observé dans l'espace ? La carte du mouvement ne serait-elle pas plutôt un cas particulier de carte de flux ? La cartographie de routes exactes des flux agrégés est délicate, car elle suppose que tous les individus définissant l'agrégat aient emprunté la route représentée, ce qui ne serait possible, à l'échelle mondiale, que dans certains cas particuliers.

La situation est différente à l'échelle locale car les différents types de flux empruntant les principaux axes de transport structurent fortement les flux. Cela se traduit par la perception d'une *distance cartographique* proche de la distance parcourue sur réseau, c'est-à-dire d'une forme de *distance géographique* stable à l'échelle locale. A cette échelle, l'application d'un angle au tracé des flux, pour le rendre courbe, n'est pas de nature à modifier le motif général de ces flux.

¹⁸³ Le procédé de C.-J. Minard (1845) sera peut-être remis au goût du jour, une rénovation est déjà accessible par les outils *Flowmap Generator*, disponibles en ligne aux adresses suivantes : <http://www.arcgis.com/home/item.html?id=3bb350838ee044979d1f583b116cb8c1> et *Distributive flow lines* <http://blogs.esri.com/esri/apl/2013/08/26/flow-map-version-2/>.

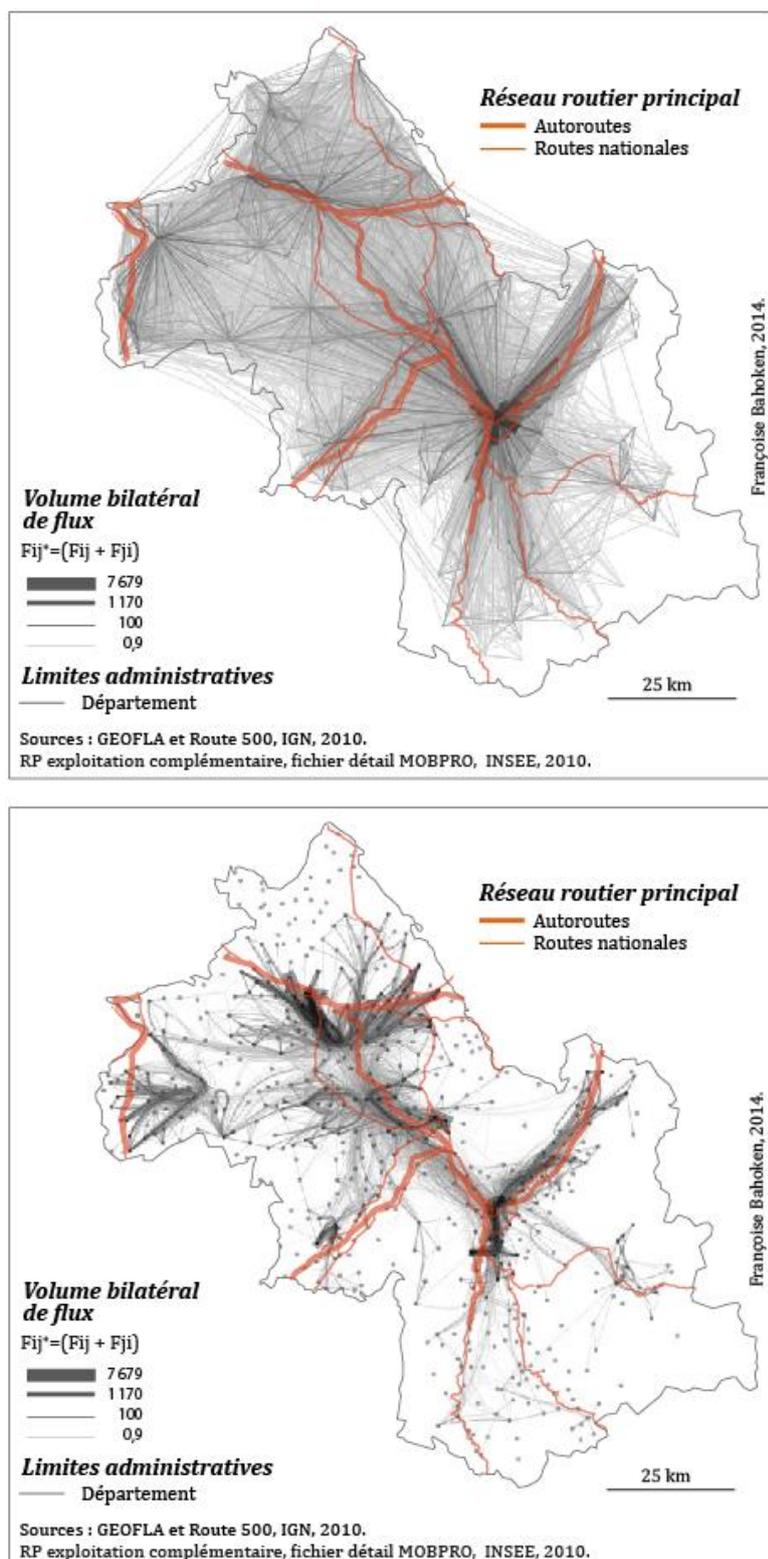
¹⁸⁴ Le tracé exact de ces routes n'étant pas accessible, leur représentation s'appuie les dires de A. Frémont, expert de la question.

Figure 6-9. Illustration de la variation de la perception des quantités échangées à l'échelle mondiale, l'exemple du trafic mondial de conteneurs (2012)



La Figure 6-10 met en évidence la concordance de la perception des différentes distances à l'échelle locale ($L_{ij} \approx D_{ij} \approx E_{ij}$), avec l'exemple du volume bilatéral (F_{ij}^+) des navettes de l'Isère. Celui-ci est sous la forme d'un trait droit sur la carte du haut et d'un trait incurvé sur la partie basse : issu d'une procédure automatique de fusion graphique des liens.

Figure 6-10. La concordance visuelle des distances cartographiques, géographiques et empiriques perçues à l'échelle locale, l'exemple du volume bilatéral (F_{ij}^+) de navettes de l'Isère



Pour la suite, nous allons donc supposer que la *distance géographique* est égale à la *distance cartographique*, toutes deux étant fixes dans un système de projection. Cela nous permet de revenir à notre troisième hypothèse qui va nous conduire à discuter la notion de **mouvement**, au regard de celle du flux, à raisonner selon une **logique de mouvement** qui prend en compte la *largeur* (Fij) ainsi que la *longueur* (Dij) du flux.

6.3 D'une logique de flux (Fij) à une logique de mouvement (Mij)

Pour corriger le décalage lié à la perception d'un **mouvement** dans la représentation spatialisée d'un flux, il convient d'agir sur la *surface* du figuré en prenant garde de ne pas modifier la *distance cartographique*, puisqu'elle est fonction de la position des lieux. Pour ce faire, il est nécessaire d'introduire une matrice de distance géographique (Dij) que l'on combinera à la matrice de flux (Fij). L'objectif de cette section est d'examiner le cas de la cartographie d'une matrice de **mouvements** $(Mij)=(Fij.Dij)$ dans le cadre de notre troisième hypothèse (voir Figure 6-1, page 249) que nous définissons de la manière suivante.

L'hypothèse n°3 suppose, en effet, que l'œil percevrait une variation de la *surface géométrique* et non seulement d'une *surface visuelle*, définie par le produit de la *largeur* (Fij) du figuré en fonction de sa *longueur* (Dij). Cette hypothèse inscrit notre raisonnement dans un espace, où l'observateur visualise le déplacement d'une quantité (Fij) correspondant, formellement, à un **mouvement (Mij)**. Supposer cette hypothèse exacte revient alors à admettre que la cartographie de flux spatialisés décrit une quantité de **mouvement**. Nous définissons cette quantité (Mij) dans la section suivante avant d'examiner les aspects liés à sa perception.

6.3.1 Définition de la quantité de mouvement (Mij)

Afin d'éviter de compliquer inutilement la présentation des concepts mobilisés dans la définition du **mouvement** que nous proposons, nous allons nous placer dans une situation idéale de corrélation parfaite ou quasi parfaite entre les trois mesures d'éloignement des lieux envisageables que sont : la *distance-temps* exprimée en minutes ; la *distance géographique* en kilomètres et la *distance cartographique* en millimètres. Cette situation idéale revient, en pratique, à supposer que :

- **Les déplacements réalisés dans l'espace se font à vitesse constante et en ligne droite**, ce qui nous permet d'assurer la proportionnalité entre la distance kilométrique parcourue dans la réalité et le temps de déplacement ;
- **La projection cartographique et l'échelle de la carte** respectent un rapport de proportionnalité strict entre la distance mesurée à la surface du globe, la *distance géographique* et la *distance cartographique* ;
- **La symbolisation des flux se fait à l'aide d'un trait droit** faute de quoi la courbure (même esthétique) entraînerait des distorsions entre la *distance cartographique* et le déplacement spatio-temporel effectivement réalisé.

Le non-respect éventuel de l'une ou de l'autre de ces conditions entraîne une modification du raisonnement que nous précisons, en cas de besoin, au cours de la démonstration.

La question qui se pose dès lors est la suivante. Quelle est la raison pour laquelle l'œil perçoit un **mouvement** $(Mij=Fij.Dij)$ sur la cartographie d'un flux (Fij) et non l'information de flux qui est à l'origine de la figure ? L'introduction d'une **logique de mouvement** nous permet d'y répondre. Elle conduit à générer une nouvelle matrice de mouvement, comme suit.

Soient deux matrices : une matrice de flux (Fij), que nous supposons symétrique, dans un premier temps, correspondant par exemple à (Fij^+) et une matrice de distance (Dij), également symétrique, par définition. Nous pouvons en déduire au moins deux autres matrices issues de leur combinaison : l'une

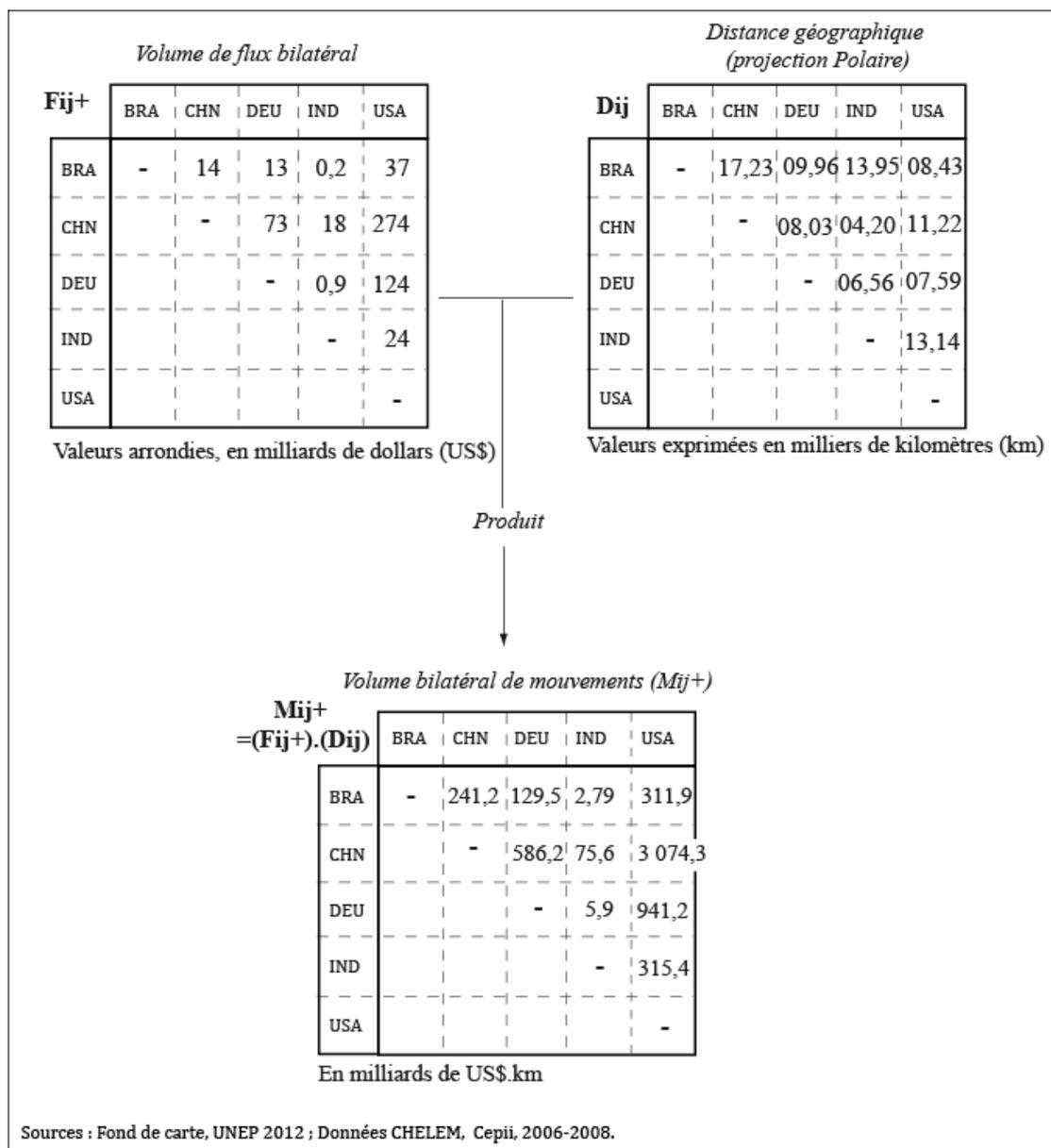
est formée par leur produit (M_{ij_1}) et l'autre (M_{ij_2}), par leur rapport, respectivement formalisées par [6.1] et [6.2] :

$$(M_{ij_1}) = [(F_{ij}).(D_{ij})] \tag{6.1}$$

$$(M_{ij_2}) = [(F_{ij}) / (D_{ij})] \tag{6.2}$$

La Figure 6-11 présente la construction de (M_{ij_1}) appliquée à notre exemple de matrice monde, en projection polaire, qui correspond à ce que nous qualifions de matrice de *mouvement* (M_{ij}).

Figure 6-11. La construction de la matrice de mouvement spatial $M_{ij_1}=(F_{ij}.D_{ij})$

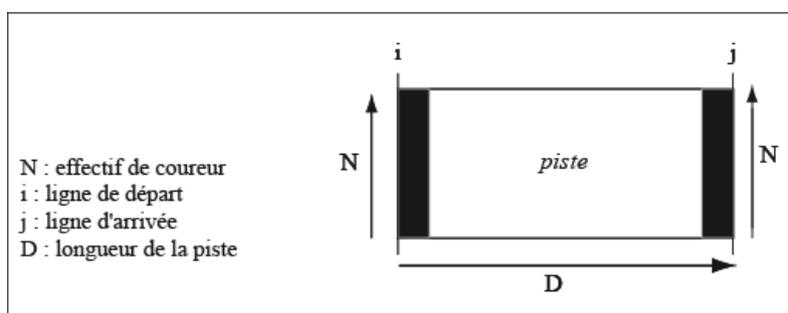


(M_{ij_1}) définit une matrice de *mouvement spatial* $M_{ij}=(F_{ij}.D_{ij})$. Le terme (M_{ij_2}), qui correspond à la matrice (I_{ij}) n'est volontairement pas représenté sur la Figure 6-11, car il n'est pas exploitable sous une forme cartographique en l'état ; nous verrons qu'elle correspond en fait à l'hypothèse n°2 que nous avons déjà examinée et éliminée faute de pouvoir la développer ici davantage.

Le *mouvement spatial* (M_{ij}) est une nouvelle quantité qui prend en compte la distance parcourue dans la définition de la valeur du flux à cartographier. Si la construction de cette nouvelle matrice ne soulève pas de difficulté particulière, son interprétation, sa dénomination et, par conséquent, sa représentation cartographique nécessitent une explication détaillée.

Une première idée intuitive et simple de la signification de ce *mouvement spatial* est celle de la mesure du débit d'un cours d'eau, ou du déplacement d'un groupe de coureurs sur une piste d'athlétisme. Pour la comprendre, imaginons la situation d'une quantité de coureurs (soit une matrice de type F_{ij}^+) qui traverse une piste de cent mètres de long ($D_{ij}=100$), reliant une ligne de départ (i) à une ligne d'arrivée (j). Ces deux lignes présentent une quantité de coureurs identique qui correspond au nombre N d'individus ($N=F_{ij}^+$), comme illustré sur la Figure 6-12.

Figure 6-12. L'analogie du mouvement avec une course à pied



La piste étant rectangulaire, les lignes de départ et d'arrivée correspondent à la *largeur* du flux. Après le signal de départ, la course s'effectue au cours d'une certaine période de temps $[t, t_{+1}[$ que nous ne connaissons pas. Si nous connaissons l'effectif (F_{ij}) de coureurs au départ et à l'arrivée, nous ignorons le moment exact du franchissement de la piste par les différents individus, car nous raisonnons sur un agrégat et la mesure est rétrospective, comme nous l'avons vu dans le chapitre 1. Nous ne savons donc pas quel est le cheminement exact des coureurs entre les deux lignes de départ et d'arrivée. Ce dont nous sommes en revanche sûrs, c'est de la quantité d'énergie « dépensée » par ces coureurs puisque chacun d'entre eux a traversé la piste.

Ainsi, si nous observons 4 coureurs au départ de la course, nous constatons que lorsqu'ils auront franchi les 100 mètres de la piste, ils auront « produit » $4 \times 100 = 400$ mètres de déplacement, en raison des possibilités d'additivité (en ligne et en colonne) des valeurs d'une matrice. De ce fait, ces 4 coureurs auraient tout aussi bien pu faire un relais que le résultat serait dans notre cas identique. Ces 400 mètres correspondent, en réalité, à l'effort physique de ces coureurs qui peut être interprété, dans notre cadre d'analyse, comme une *quantité de mouvement*. Nous faisons ici **l'analogie entre la dépense physique d'énergie de ces coureurs et le frottement physique lié au déplacement dans l'espace, équivalent au frein de la distance géographique** déjà mentionné. De ce fait, cette quantité correspond à une mesure de *mouvement spatial* et non de changement dans l'espace, comme celle de flux.

La *quantité de mouvement* produite par les coureurs, lorsqu'ils franchissent la piste, correspond à un stock « d'énergie » affecté aux différentes cases de la matrice : la valeur (i,j) de la matrice (M_{ij}^+) présentée sur la Figure 6-11, page 273 est bien un effectif de population et c'est parce que cet effectif est transféré dans l'espace, que l'espace est pris en compte dans la mesure, qu'il s'agit d'un mouvement. Cette matrice conserve ses propriétés issues du tableau de contingence : le stock dont il est question peut être additionné sur tout ou partie de la matrice ; elle conserve aussi sa propriété

générale de symétrie comme définie dans le chapitre 1. Ce stock est en outre fini, puisque les données de flux observées sont discrètes, tout comme les distances parcourues par ces flux.

La *quantité de mouvement* peut être exprimée dans l'unité de mesure de la métrique d'éloignement choisie (dans notre exemple, un nombre de coureurs.kilomètres) ce qui ouvre la voie vers de nombreuses applications théoriques, empiriques et cartographiques que ne permet pas le flux. Par exemple, dans le cas d'une matrice de flux de marchandises réalisés par voie routière, la somme totale des mouvements (de chaque case de la matrice) définit indirectement la quantité d'énergie dépensée par les véhicules, c'est-à-dire la quantité de carburant consommée mais aussi leur empreinte environnementale, à travers la mesure de leur bilan carbone. De la même façon que les flux internationaux de passagers rapportés au kilométrage parcouru peuvent être exprimés en termes de bilan carbone liés à la distance parcourue et au mode aérien, les trajets en train, sur un réseau ferroviaire, peuvent être aussi exprimés en termes de bilan carbone,

La quantité de mouvement étant définie, examinons sa représentation cartographique au regard de celle du flux, dans l'hypothèse **d'une perception du mouvement sur la cartographie d'une matrice de flux**.

6.3.2 L'hypothèse n°3 d'une perception du 'mouvement'

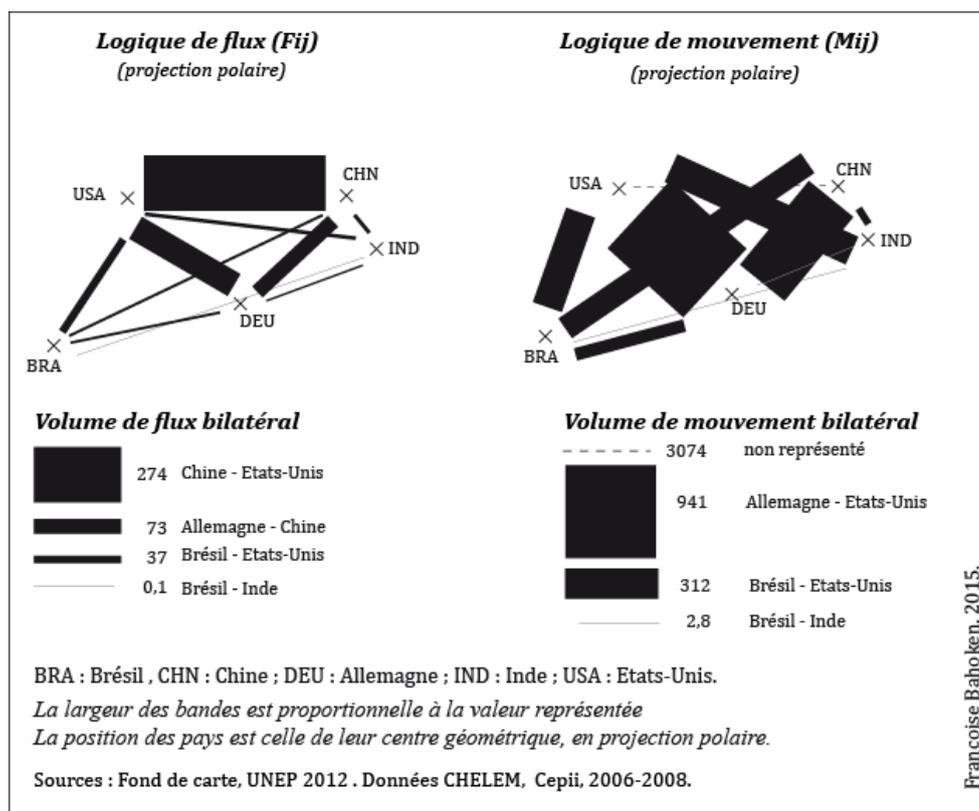
L'hypothèse d'une perception du mouvement (M_{ij}) est celle de la perception d'une tâche en contraste d'un fond (voir Figure 6-1, page 249), de la représentation de (F_{ij}) en fonction de (D_{ij}) correspondant bien à la quantité de mouvement (M_{ij}) que nous venons de définir.

Cartographier ce mouvement conduit à une correction de la *largeur* du figuré de sorte que sa surface réelle corresponde à la surface perçue, au mouvement, la *longueur* étant nécessairement fixée dans un système de projection. Cette correction de la *surface* se manifeste en réalité par une diminution de la *largeur* des figurés qui parcourent de faibles distances et par une augmentation de celle de ceux qui parcourent de longues distances. C'est d'ailleurs ce processus qui, en se renforçant lorsque la valeur de flux augmente, peut empêcher¹⁸⁵ la représentation directe, de manière proportionnelle, de la valeur (M_{ij}^+) pour USA-CHN sur la carte suivante. La Figure 6-13 illustre la différence entre la réalisation d'une carte de flux selon une *logique de flux* (F_{ij}) (placée à gauche) et selon une *logique de mouvement* (M_{ij}).

On observe ainsi que la *logique de mouvement* conduit à survaloriser la perception des flux les plus longs – donc qui parcourent les distances les plus importantes (par exemple : BRA-DEU ou encore DEU-CHN) –, qui sont aussi généralement les plus petits – en termes de largeur, ils présentent logiquement les valeurs les plus faibles – et à sous-estimer les autres (sauf dans certains cas particuliers tels que USA-CHN). En effet, d'après la Figure 6-13, l'adoption d'une *logique de mouvement* dans l'hypothèse n°3 entraîne le doublement de la perception de l'espace, puisqu'il est représenté deux fois. En effet, la *largeur* du figuré étant égale à (M_{ij}), sa *surface* est alors égale à $S=(F_{ij}).(D_{ij})^2$ avec une quantité de noir de 100%. Cette double représentation de la distance est liée à sa présence dans la formulation du flux et dans sa représentation.

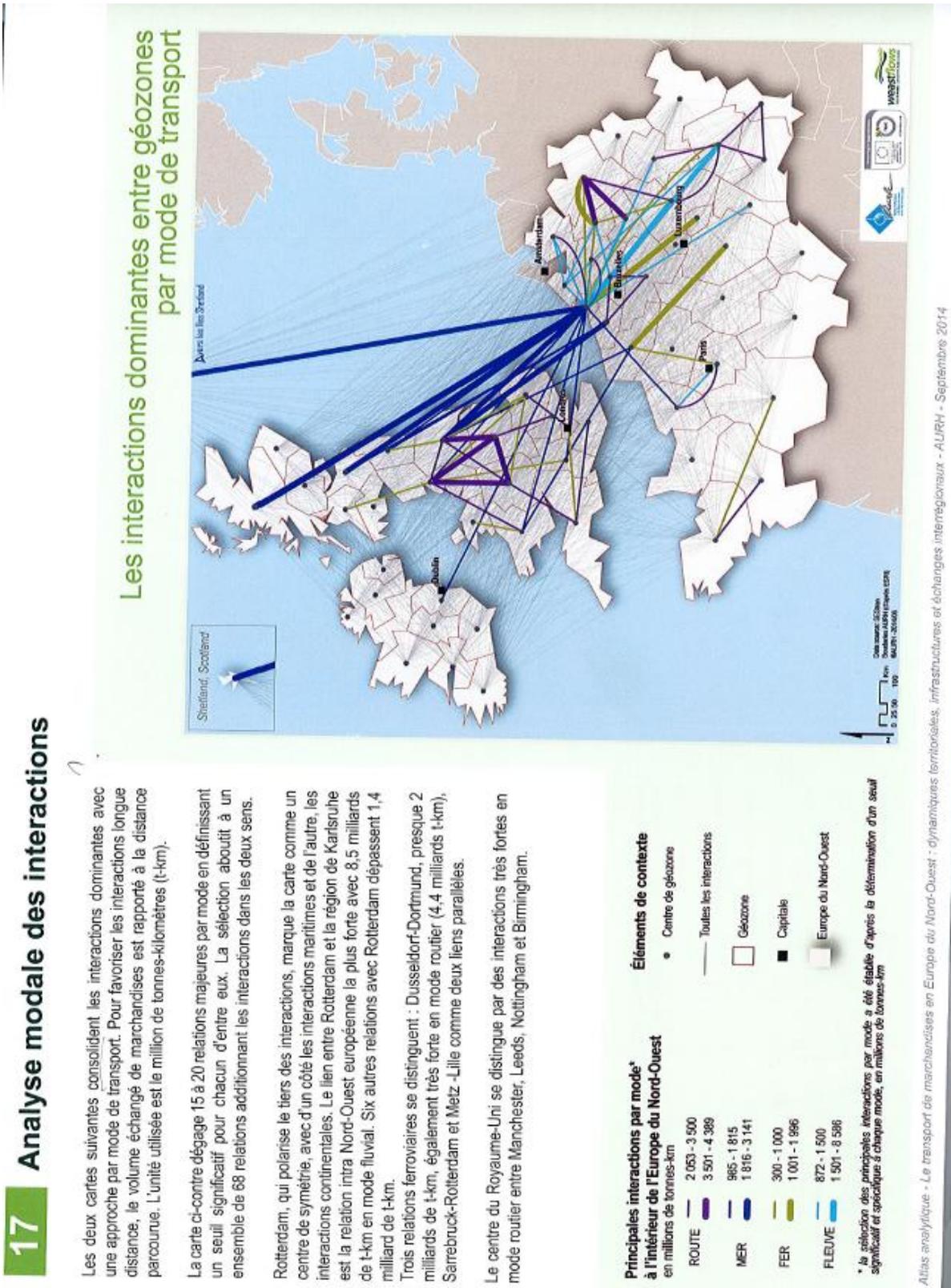
Dit autrement avec un exemple, sur une carte de flux de marchandises exprimées en tonnes.kilomètres, on pourrait imaginer les cartographier de sorte que la *largeur* des figurés soit proportionnelle au mouvement (M_{ij}). Cependant, comme $(M_{ij})=(F_{ij}).(D_{ij})$, procéder ainsi reviendrait à faire intervenir deux fois la distance : à la fois dans la *largeur* du figuré et dans sa *longueur* (puisque l'œil perçoit cette dimension).

¹⁸⁵ Il est toutefois possible d'utiliser d'autres méthodes de représentation cartographique pour contourner ce problème, mais ce n'est pas l'objet. De plus, la non représentation de ce flux n'est pas de nature à gêner la démonstration.

Figure 6-13. Illustration de la cartographie d'une *logique de flux* versus de *mouvement*

C'est la méthode mise en œuvre dans les « cartes d'interactions » de marchandises de l'Atlas analytique de l'Agence d'Urbanisme de la Région du Havre et de l'Estuaire de la Seine (AURH), telles que celle de la Figure 6-14 (page suivante). « Les interactions sont d'abord exprimées en tonnes. Elles sont aussi exprimées en tonnes-kilomètres en les multipliant par les distances théoriques de centre de zone à centre de zone. Cette unité favorise les marchandises pondéreuses » (AURH, 2014 : 54). Plus loin, il est indiqué ceci. « Pour favoriser les interactions longue distance, le volume échangé de marchandises est rapporté à la distance parcourue, l'unité est le million de tonnes-kilomètres » (AURH, 2014 : 60). Cette survalorisation des mouvements à longue distance conduit à une carte erronée pour deux raisons. La première tient au fait que les cartes doivent être interprétées en termes de mouvement et non de flux (ni, d'ailleurs, à proprement parler d'interactions, comme il est indiqué en légende) : sur la Figure 6-14, la *largeur* de la bande est fonction de la classe de valeur du flux de marchandises exprimée en tonnes.kilométriques, c'est-à-dire au mouvement et non à un flux. La seconde raison tient à l'intensité de teinte (ici 100 %) qui devrait être atténuée (voir Figure 6-1, page 249) de manière à affaiblir la perception des distances, nous reviendrons sur cette proposition dans le chapitre 7 (page 288). La mise en œuvre de la *Valeur* (100% de teinte) survalorise les flux à longue distance, en raison de l'augmentation trop importante de leur surface visuelle (de la quantité de teinte perçue). Ces longs flux qui correspondent aux figurés les plus fins étant mieux perçus, on peut dire que la méthode agit visuellement à la manière d'un critère de segmentation de la matrice qui aurait été appliqué en amont de la représentation, mais cela n'est pas le cas, comme nous le verrons dans le chapitre 7.

Figure 6-14. Exemple de carte de flux de marchandises, exprimée en tonnes kilométriques



Source : AURH, 2014.

Ce résultat, qui peut apparaître surprenant au premier abord, n'est pas forcément paradoxal puisque, d'après W. Tobler (1970b), l'objectif d'une carte de flux est bien d'illustrer les fréquences de relations élevées, c'est-à-dire les flux qui concernent dans l'ensemble beaucoup d'individus et un espace plus important que ceux qui sont moins nombreux, moins fréquents, mais présentent des valeurs élevées. Ces longs flux concernent effectivement une population plus importante que celle des courts flux, ceux qui sont importants en valeur mais qui ne parcourent pas nécessairement de longues distances (puisque l'on sait que la relation entre (F_{ij}) et (D_{ij}) est inversement proportionnelle : pour un même couple de lieux, quand (F_{ij}) croît, (D_{ij}) décroît). Cependant, retenir cette solution conduit à augmenter *l'effet de couverture spatiale* puisqu'elle survalorise la perception de la masse de petits flux diffus, de ceux-là qui sont habituellement éliminés de l'analyse parce qu'ils présentent des valeurs faibles n'étant généralement pas considérées comme *significatives*, nous y reviendrons dans le chapitre 8.

Cette double intégration de l'espace apparaît dès lors redondante, elle est aussi inutilement source de complexité graphique. Nous signalons toutefois qu'il est possible d'améliorer le résultat cartographique obtenu sur la Figure 6-13 de deux manières. La première possibilité consiste à représenter le mouvement de manière graphique : en suivant le cheminement exact emprunté par le flux, si l'information est connue, c'est-à-dire à représenter un flux affecté sur réseau. La seconde vise à atténuer l'intensité de la *surface* visuelle du figuré en utilisant soit la *Transparence*, soit la matrice $(M_{ij})_2$ évoquée page 273, ouvrant alors la voie vers la cartographie d'une intensité (I_{ij}) de mouvement qui mobilise en outre la *Valeur*, nous y reviendrons dans le chapitre 7 (page 288).

Nous pensons *in fine* que **l'hypothèse n°3** n'apporte pas grand-chose, sauf à une surreprésentation abusive des différences de taille des liens qui pourrait être préjudiciable à l'analyse. Elle est, en tous les cas, à manier avec précaution. Ainsi, tout comme **l'hypothèse n°2**, **l'hypothèse n°3 conduit aussi à une représentation erronée de la réalité, par un excès inverse : elle survalorise les flux les plus fréquents, au détriment des plus importants (en privilégiant la perception des longues distances parcourues).**

Cartographier un mouvement (M_{ij}) consiste en réalité à représenter une quantité de flux (F_{ij}) ayant parcouru une distance réelle, entre des lieux inscrits dans un référentiel. L'œil percevant naturellement une surface, la cartographie du mouvement $(M_{ij}) = [(F_{ij}/D_{ij}).F_{ij}]$, correspond à *peu près* à (F_{ij}) , à *l'information disponible* dans la matrice de départ.

Ce résultat n'est pas anodin : il est majeur et justifie l'énoncé d'une quatrième hypothèse selon laquelle l'œil percevrait, en réalité, une forme de mouvement (M_{ij}) qui correspondrait à la représentation approximée d'un flux (F_{ij}) dans un espace.

6.3.3 *L'hypothèse n°4 d'une perception 'approximative du mouvement' sur une carte de flux*

Cette quatrième hypothèse inscrit la distance parcourue, en tant que forme simplifiée de séparation spatiale, comme composante intrinsèque de la perception des flux et cela, quelle que soit l'information représentée : un **flux (F_{ij})** ou un **mouvement (M_{ij})** . Elle intègre par essence le temps puisque, selon les cas, la carte représentera un nombre d'individus qui s'est déplacé (donc un flux) ou bien un nombre d'individus par kilomètre qui s'est déplacé sur un espace géographique (donc un mouvement).

Il est important de noter que cette hypothèse n'est exacte que dans le cas où le flux est spatialisé, c'est-à-dire représenté sur une carte et non sur un graphe ordinaire. Elle correspond aussi, sur le plan de la construction à l'hypothèse n°1, ce qui signifie que **les différentes cartes réalisées sous les hypothèses n°1 et n°4 sont rigoureusement identiques sur le plan de la construction cartographique, mais elles diffèrent sur le plan sémantique, en fonction de la manière dont la distance est appréhendée, c'est-à-dire en fonction de l'interprétation donnée au flux** (voir Figure 6-16).

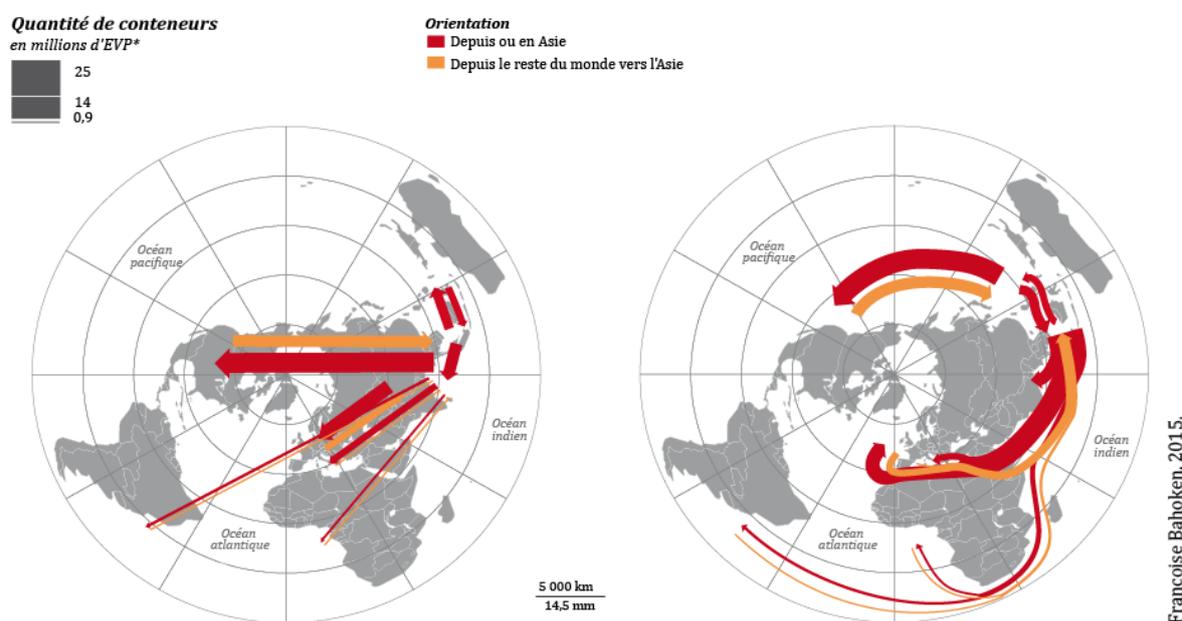
En reprenant l'analogie des coureurs, si nous supposons toujours que la *distance cartographique* correspond à la *distance euclidienne* ($L_{ij} = D_{ij}$), dans un contexte où la projection cartographique de la sphère sur le plan est importante, quelle que soit l'échelle de représentation, le fait que la perception des flux soit approximativement celle d'un mouvement ($F_{ij} \approx M_{ij}$) signifie que l'observateur statique d'une course à pied perçoit le déplacement d'une *largeur* dans un *espace* qui correspond bien à la *surface* de la bande calculée comme le produit de sa *longueur* * sa *largeur*.

Si, comme nous l'avons dit, à l'échelle locale, ce flux peut être observé sur un réseau et la *distance cartographique* correspondre à peu près à la *distance fonctionnelle*, alors la cartographie de ce flux coïncide bien avec la représentation de la quantité d'individus mobiles au cours du temps sur un espace, c'est-à-dire avec (M_{ij}). Pour que ces deux quantités soient égales, il faudrait que la représentation du flux suive effectivement un réseau de transport faute de quoi, la *distance cartographique* ne saurait restituer à elle seule la *distance fonctionnelle* et ($F_{ij} \approx M_{ij}$).

Si les cartographies de (F_{ij}) et de (M_{ij}) sont identiques à l'échelle locale, dans leur construction, qui correspond à l'hypothèse n°1, elles le sont aussi, plus généralement, sur le plan de la perception. Ce qui change en réalité entre ces deux quantités c'est avant tout l'interprétation que l'analyste réalise des flux qui sont ainsi cartographiés. **Puisque la construction d'une carte de flux est invariante, ce qui changera entre deux cartes de flux, c'est par conséquent leur sémantique** (voir Figure 6-15).

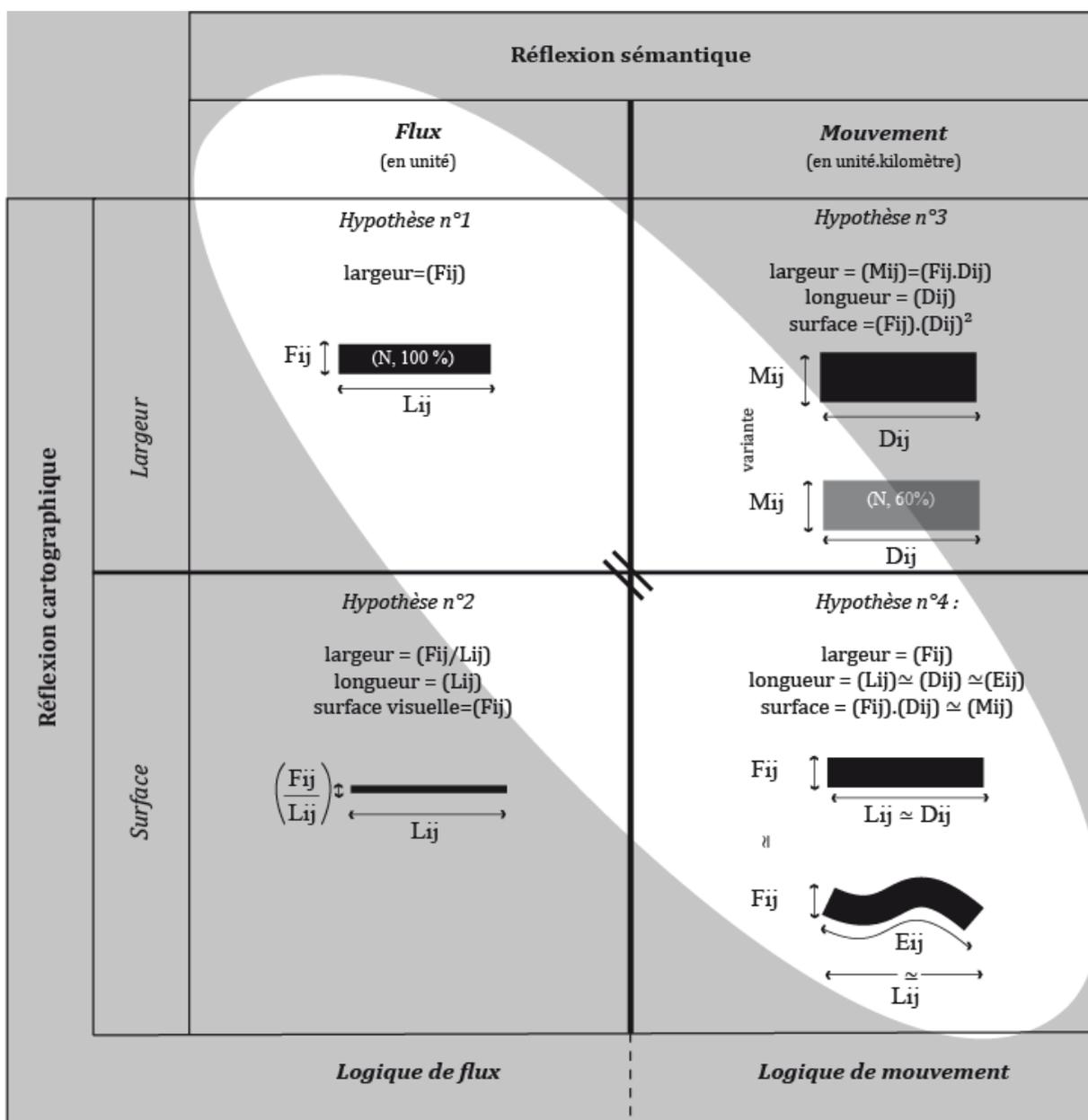
En effet, sur le plan de la construction cartographique, quelle est la différence entre les deux cartes suivantes déjà présentées, hormis le fait que l'une décrit un flux de marchandises dans le monde et l'autre, un mouvement mondial de marchandises par voie maritime ?

Figure 6-15. Cartographie de flux versus cartographie de mouvement



La cartographie de flux est bien, en réalité, une approximation de la représentation du mouvement, car la distance cartographique – lorsqu'elle est considérée dans le formalisme de la métrique euclidienne – ne saurait correspondre, avec exactitude, à la distance fonctionnelle. On voit bien sur la Figure 6-15, la différence de perception des distances parcourues selon que la métrique est euclidienne ou fonctionnelle, que le tracé est linéaire ou sinueux. C'est pourquoi les hypothèses n°1 et 4 sont identiques, sur le plan de la construction et diffèrent sur le plan de leur interprétation (voir Figure 6-16).

Figure 6-16. Cadre sémiotique de réflexion sur la construction de la carte de flux (complété)



Dimensions du figuré linéaire

Largeur

- Fij : quantité transférée d'un lieu à un autre, dans un plan
- Mij : quantité transférée d'un lieu à un autre pondérée par la distance parcourue

Longueur :

- Lij : distance cartographique perçue (= longueur du figuré)
- Dij : distance géographique, euclidienne ou réseau (= longueur du figuré)
- Eij : distance empirique ou coût généralisé de transport (= longueur du figuré)

Surface :

- surface visuelle (N) : quantité de teinte noire, exprimée en pourcentage (%) de CMJN
- surface géométrique (S) : produit de la longueur par la largeur du figuré

Françoise Bahoken, 2015.

Ce résultat souligne l'importance de la sémantique (flux *versus* mouvement), aussi celle de la thématique dans l'interprétation d'une carte de flux (flux de migrants *versus* flux de migrations *versus* mouvements migratoires / flux financiers *versus* mouvements financiers / flux de marchandises exprimés en tonnes.kilomètres ou en valeur monétaire..., que nous avons déjà souligné dans le chapitre 1, lors de la mesure du flux.

On se rend maintenant mieux compte de l'impact de la cartographie de flux réalisée à l'échelle mondiale, de ces flux de marchandises exprimés en tonnage *versus* mouvements financiers de biens dont le choix conduit à un arbitrage entre la représentation des routes empruntées par les valeurs monétaires ou non, ces dernières correspondant au financement des marchandises dont il est question. Ce choix est (hautement) délicat car il peut conduire, en fonction de la thématique du flux, à mettre en évidence le trajet emprunté par certains mouvements financiers dont on se rendra compte qu'ils ne sont pas nécessairement il(lé)gaux, ou bien qu'ils ne suivent pas le trajet le plus direct qui existe entre deux lieux ! De la même façon, on peut s'interroger sur la nécessité de représenter les mouvements migratoires à la place des flux de migrants, devant conduire à identifier le trajet qu'ils empruntent, les routes par exemple pour rejoindre l'Europe du nord depuis les pays situés dans son voisinage immédiat, dont on sait pertinemment qu'ils ne sont pas forcément (il)légaux ?!.

Conclusion du chapitre

Ce chapitre a permis de montrer qu'il n'est pas possible d'agir sur la *distance cartographique* sans changer la perception de l'information de flux représentée sur la *largeur* du figuré, au droit des lieux d'origine et de destination. Cela signifie d'une part, que le mode de construction de la carte de flux est invariant, c'est-à-dire **qu'il n'y a qu'une et une seule solution de cartographie qui correspond aux hypothèses n°1 et n°4** et, d'autre part, que ce qui compte avant tout dans la réalisation d'une carte de flux, c'est sa *sémantique*, c'est-à-dire la thématique dont il est question et l'interprétation que l'analyste (et/ou l'auteur de la carte) en fait.

Il y a, en effet, deux interprétations à la carte de flux classique, qui correspondent aux deux *logiques de flux et de mouvement* sous-jacentes à leur représentation. Dans le premier cas, qui correspond à l'hypothèse n°1, seule la largeur du figuré importe dans la représentation de la matrice de flux dans le plan et à la limite, il n'est pas nécessaire de l'assortir d'une métrique, ce qui pose la question de l'utilité de la carte de flux au regard des autres types de graphiques (du graphe, par exemple). En revanche, dans le second cas, la représentation du mouvement dans le plan nécessite le recours à une acception de la distance parcourue par le flux, *a fortiori* à son inscription dans un espace assorti d'une métrique. Elle correspond à l'hypothèse n°4 selon laquelle la carte de flux correspond à une représentation approximée du mouvement dans le plan.

Les hypothèses ayant résisté à l'investigation menée au cours de ce chapitre sont la première et la dernière, les hypothèses n°2 et n°3 correspondent en fait à des fausses pistes, sauf à recourir à la *Valeur*, en mobilisant une forme d'intensité de mouvement ou, sur le plan sémiologique, en recourant à la *Transparence*. Si l'hypothèse n°1 correspond à l'approche classique, l'hypothèse n°4 lui est similaire sur le plan de la construction. Ainsi, ce qui change entre ces deux hypothèses, entre une carte de flux (hypothèse n°1) et une carte de mouvement (hypothèse n°4), ou bien d'une carte de flux à une autre, ce sont les éléments qui sont portés en légende, c'est-à-dire la manière dont l'information cartographiée (les lieux, les valeurs, ...) est sélectionnée.

Lorsque la matrice est dense, la complexité graphique de la carte qui en résulte nécessite de sélectionner l'information à représenter, dans le cas où il n'est pas question de représenter *toute l'information disponible* dans la matrice comme suggéré dans le chapitre 5. La question qui se pose alors est celle de savoir comment on sélectionne l'information à cartographier : sur le plan sémantique, selon une *logique de flux* ou bien selon une *logique de mouvement estimé*, par l'application d'un angle au tracé du lien, par exemple. Nous avons vu qu'à l'échelle locale, la perception des trois distances était équivalente ($L_{ij} \simeq (D_{ij}) \simeq (E_{ij})$). La cartographie qui en découle mobilise alors une distance sur réseau, ce qui prouve bien que la cartographie des flux immatériels ne saurait être qu'une approximation de celle du *mouvement*. De ce point de vue, la solution qui consiste à réaliser des figurés courbes plutôt que droits correspond à une pratique certes esthétique, mais dont l'objectif premier est de donner l'illusion de la représentation d'un mouvement, de représenter une approximation du mouvement. E. G. Ravenstein ne s'y était d'ailleurs pas trompé en 1885 en représentant de cette manière, la multitude de petites flèches curvilignes, afin de donner l'apparence d'un mouvement continu localement. C'est aussi l'un des éléments d'explication de l'origine des méthodes automatisées de fusion graphique des liens qui ont été développées afin de mieux visualiser les mouvements locaux.

Quoi qu'il en soit, **la question qui se pose maintenant est celle de la signification de la cartographie des n flux sélectionnés, car elle est susceptible de varier en fonction de la logique (de flux ou de mouvement) sous-jacente, mais aussi de la manière dont le ou les critères de sélection de ces n flux ont été définis, en amont de l'étape de cartographie proprement dite**. En effet, la cartographie des 100 plus gros flux variera selon que l'on s'intéresse aux 100 premiers flux ou aux 100 premiers mouvements.

C'est pourquoi les trois prochains chapitres seront consacrés au choix du critère de sélection dont on verra qu'il pourra être déterminé selon deux grandes approches : en ayant recours à une matrice externe qui servira de variable de restriction aux valeurs de (F_{ij}) ou en appliquant directement un ou plusieurs critères de sélection directement sur la matrice (F_{ij}) .

Ainsi, dans le chapitre 7, la matrice (F_{ij}) sera nuancée en fonction d'une matrice d'éloignement, de voisinage ou d'appartenance régionale des lieux tandis que dans les chapitres 8 et 9, un critère de sélection sera appliqué, respectivement au niveau global de l'ensemble de la matrice ou local des lieux d'origine et/ou de destination.

7^{ème} CHAPITRE. Propositions de solutions liées à la prise en compte de l'espace : la distance, le voisinage spatial et l'appartenance territoriale

« J'évoque la première loi de la géographie : toutes les choses sont reliées entre elles, mais celles qui sont proches le sont davantage que celles qui sont éloignées. »¹⁸⁶

Waldo Rudolf Tobler (1970).

Introduction du chapitre

En raison de multiples contraintes physiques entre les lieux qui agissent sur la vitesse de déplacement, *a fortiori* sur leur éloignement, la cartographie d'une matrice de flux est susceptible de varier en fonction de l'acceptation de la distance, comme nous l'avons vu dans le chapitre 6, mais aussi en fonction de l'échelle géographique d'observation des mobilités, c'est-à-dire selon qu'elle s'exprime à *courte distance* ou bien à *longue distance*.

Nous avons montré dans le chapitre 6 l'importance de la représentation de la *distance cartographique* (Lij), traduisant la perception de la *distance géographique* (Dij) parcourue par le flux. Si la *distance euclidienne* traduit la valeur minimale de la distance géographique, il est également possible de raisonner en termes de minimum kilométrique parcouru sur un réseau, de minimum temporel, financier ou de minimum perçu (par exemple, en fonction du temps passé dans un mode de transport lors d'un déplacement entre deux lieux). Le choix de la mesure de *distance géographique* à mobiliser est par ailleurs d'autant plus sensible que l'échelle est globale. Ainsi, lorsque l'espace concerné est projeté, les premières variables d'intérêt sont la manière dont l'information de flux (Fij) est appréhendée et l'interprétation qui en est faite, conduisant à la qualification de la figure de carte de flux spatialisé (Fij) ou de carte de mouvement (Mij=Fij.Dij) par opposition au graphe ordinaire. Or, le chapitre 6 a également mis en évidence la possibilité de percevoir plusieurs classes de distances (Dij), dans l'éventualité où (Lij) et (Dij) coïncident à peu près, ce qui correspond à plusieurs échelles de la mobilité.

Sur le plan cognitif, la prise en compte de l'éloignement des lieux dans le processus cartographique intéresse le sens attribué par un observateur à l'information de flux qui est représentée, à s'interroger, par exemple, s'il est plus important de visualiser le déplacement de 1000 personnes sur 10 km ou celui de 400 personnes déplacées sur 10 000 km ; de décrire les flux qui s'expriment à moins de 100 km ou bien à plus de 10 000 kilomètres.

Malgré leur efficacité, les améliorations sémiologiques et esthétiques proposées dans le chapitre 5 peuvent apparaître insuffisantes pour décrire des motifs spatiaux spécifiques aux flux ou aux mouvements, par exemple à la courte distance, car elles ne conduisent pas à une sélection réelle de l'information représentée. A cette échelle, la composante spatiale génère des problèmes spécifiques de

¹⁸⁶ Texte traduit par nos soins, la formule originale est la suivante : « *I invoke the first law of geography: everything is related to everything else, but near things are more related than distant things* ».

lisibilité liés à la finesse de la résolution spatiale (Bahoken, Olteanu-Raimond, 2013)¹⁸⁷. Pour les résoudre, une famille de solutions consiste à raisonner sur la situation des lieux, c'est-à-dire du point de vue de l'espacement relatif des lieux les uns par rapport aux autres en mobilisant l'espace géographique. Pour cela, trois grandes approches sont envisageables.

La première consiste à transformer l'espace géographique, en modifiant la position relative des lieux d'origine et de destination. C'est l'enjeu des changements de projection cartographique, des déformations morphologiques comme dans le cas des anamorphoses ou de leurs positions, comme dans le cas de CARTOD¹⁸⁸ (Bonin, Bahoken, Mermet, 2013).

La seconde approche consiste à déformer la géométrie des liens à l'aide de solutions qui sont graphiques, esthétiques telles que l'application d'un angle de type courbes de Bézier, d'une fusion graphique (Holten, 2005) ou mathématique (Tobler, 1969), conduisant par exemple aux cartes chiffonnées (L'Hostis, 1993).

La troisième direction consiste non pas à transformer ou à déformer la géographie des lieux et/ou des liens mais à la prendre en compte dans le processus de cartographie de flux. Cela passe par le recours à la notion d'éloignement ou de proximité des lieux qui, tout en conservant la situation des lieux et l'aspect rectilinéaire des liens, conduit à une sélection de l'information en fonction d'un critère d'éloignement (ou de proximité) des lieux.

Ce septième chapitre a ainsi pour objectif de présenter des solutions liées à la prise en compte de l'éloignement des lieux lors du traitement cartographique (descriptif) des données de flux à cartographier et non lors de leur modélisation. Il va contribuer à résoudre les problèmes spécifiques liés aux *effets de profondeur* et *d'itinéraires*, mais aussi à l'*effet de couverture spatiale*.

Nous mobilisons une matrice complète et symétrique, qui décrit un volume bilatéral de flux (F_{ij}^+). L'objectif est de nuancer cette matrice, d'en faire une matrice creuse, en recourant à une matrice externe qui traduit l'éloignement ou la proximité entre les lieux, dans le cadre d'une opération que nous notons ($F_{ij} \leftarrow D_{ij}$).

L'approche proposée ici est réalisée en dehors du processus de cartographie proprement dit, car elle agit d'abord sur (D_{ij}) pour produire un résultat qui filtrera la sélection des éléments de (F_{ij}) jugés les plus utiles à l'analyse. Elle ne conduit pas à générer une nouvelle quantité, ou un nouvel indicateur comme dans le cas du chapitre 6.

Sur le plan formel, en considérant être dans une situation de corrélation quasi parfaite entre les trois formes de distances qui peuvent être mobilisées ($E_{ij} \simeq D_{ij} \simeq L_{ij}$), il est possible de mobiliser l'éloignement des lieux de trois manières, que nous présentons tour à tour dans les sections suivantes.

Dans la première section, nous recourons à la distance géographique parcourue par le flux (D_{ij}), comme mesure de la distance réelle ($F_{ij} \leftarrow D_{ij}$). Cette mesure étant quantitative continue, nous allons la discrétiser en k classes de distances (D_{ij}) parcourues par les flux à partir desquelles nous segmenterons la matrice de flux (F_{ij}). Nous proposons deux raisonnements : le premier est d'ordre sémiologique et mobilise la *Transparence* introduite dans le chapitre 5 et le second est statistique, il consiste à transférer les méthodes de discrétisation cartographique des données de stock à celle des

¹⁸⁷ Nous avons pu montrer le rôle joué par la taille de la maille géographique sur l'information de flux dont on dispose, la déperdition de l'information à mesure que la résolution augmente, en reconstituant des matrices de flux à différentes échelles spatiales et temporelles, à partir de trajectoires individuelles de mobiles (Bahoken, Olteanu-Raimond, 2013).

¹⁸⁸ L'un des objectifs de CARTOD est de représenter les flux sur un diagramme de chord plutôt que sur une carte. Cette solution étant liée aux transformations cartographiques, nous n'allons pas la développer car elle entraîne une modification substantielle de notre cadre d'analyse à savoir un changement de la position des lieux.

flux. Dans les deux cas, la procédure conduit à la réalisation d'une collection de k cartes : $(F_{ij}) = (F_{ij}, D_{ij_1}) + (F_{ij}, D_{ij_2}) + \dots + (F_{ij}, D_{ij_k})$.

Dans la seconde section, la mesure suggérée est celle d'une proximité discrète entre les lieux. Elle correspond à la prise en compte d'un critère d'éloignement qui met en évidence un *effet de voisinage*. Nous considérons pour cela une mesure d'éloignement discrète qui sera appréhendée, dans un premier temps, de manière binaire (C_{ij}) puis, dans un second temps, sous une forme ordinale (CK_{ij}). La mesure binaire conduit à ne retenir que les flux réalisés entre des lieux proches, c'est-à-dire qui partagent une limite de zone (ou une frontière) commune. Elle mobilise une matrice de *contiguïté* (C_{ij}) d'ordre 1 qui deviendra plus tard ordinale (CK_{ij}) où le nombre k exprime une distance topologique exprimée en nombre de frontières qu'il convient de franchir pour aller d'un point à un autre.

La troisième section diffère des deux précédentes en ce sens où la proximité (ou l'éloignement) des lieux est appréhendée au regard de leur appartenance à une même entité spatiale (ou plutôt territoriale), à une même « région ». En mobilisant la notion d'appartenance territoriale, deux lieux sont considérés comme proches s'ils appartiennent à la même unité spatiale de niveau supérieur ($F_{ij} \leftarrow [P_i = P_j]$). L'appartenance considérée est politique en raison du recours à des limites de zones d'origine administrative, mais elle aurait pu être culturelle, linguistique, historique... La proposition consiste alors à ne cartographier que les valeurs de flux qui concernent des lieux appartenant à une même région ou à deux régions différentes.

Enfin, la dernière section a pour objectif de valider les solutions proposées sur une matrice de flux internationaux de passagers aériens inters aéroports (OACI, 2000).

7.1 La métrique réelle ou *distance géographique parcourue par le flux* (D_{ij})

La sélection des flux selon un critère de distance réelle mobilise une mesure d'éloignement continue, la *distance empirique* (E_{ij}), ou la *distance géographique* (D_{ij}) parcourue par le flux.

En présence d'un *effet-spaghetti* maximal et puisqu'il n'est pas forcément nécessaire de représenter toute l'*information disponible* dans une matrice de flux, deux approches utilisant la *distance géographique* parcourue par le flux peuvent être mises en œuvre. La première consiste à sélectionner les flux (F_{ij}) à représenter, en fonction d'une valeur précise de distance (D_{ij}) et la seconde à réaliser une collection de cartes qui segmente l'information disponible en différentes catégories de distances. Une manière de représenter la distance-temps passe en effet par la collection de carte (ou type bande-dessinée), un procédé dont nous verrons qu'il a été inventé par un médecin, J.-E. Marey ((1830-1904), pour la représentation du mouvement dans le plan.

Dans le premier cas, il s'agit de ne représenter, par exemple, que les flux qui parcourent moins de 100 km, qui correspondent à un type de mobilités locales. Si cette solution présente un intérêt certain lorsqu'il s'agit de représenter certaines catégories de flux, elle pose aussi le problème de la détermination d'un critère unique de sélection (ou d'élimination) des flux, sur lequel nous reviendrons dans le chapitre 8. Nous allons plutôt examiner la seconde famille de méthodes qui considèrent l'ensemble de l'*information disponible* dans une matrice de flux. Pour cela, au lieu de sélectionner des valeurs de flux à représenter, la méthode segmente plutôt la matrice de flux selon des classes de distance parcourue de manière à représenter toute l'information.

Il existe deux manières de représenter l'*information disponible* dans une matrice de flux. La première est de nature sémiologique, elle consiste à n'agir que sur l'apparence visuelle des figurés, sans toucher aux données numériques. La seconde est liée aux données, à leur construction cartographique. Elle reprend l'idée de « *collection de cartes à 1 caractère* » chère à J. Bertin pour représenter de manière simple une information complexe, en réponse à la question « *tel caractère, où est-il ?* ». Ces collections sont souvent temporelles, plus rarement catégorielles ; un exemple récent de collection de cartes de flux temporelles est disponible dans (Richard, Zanin 2009 : 144). Dans notre cas, la

collection de cartes n'est ni temporelle ni catégorielle. Elle est réalisée en fonction de classes de distances parcourues par les flux.

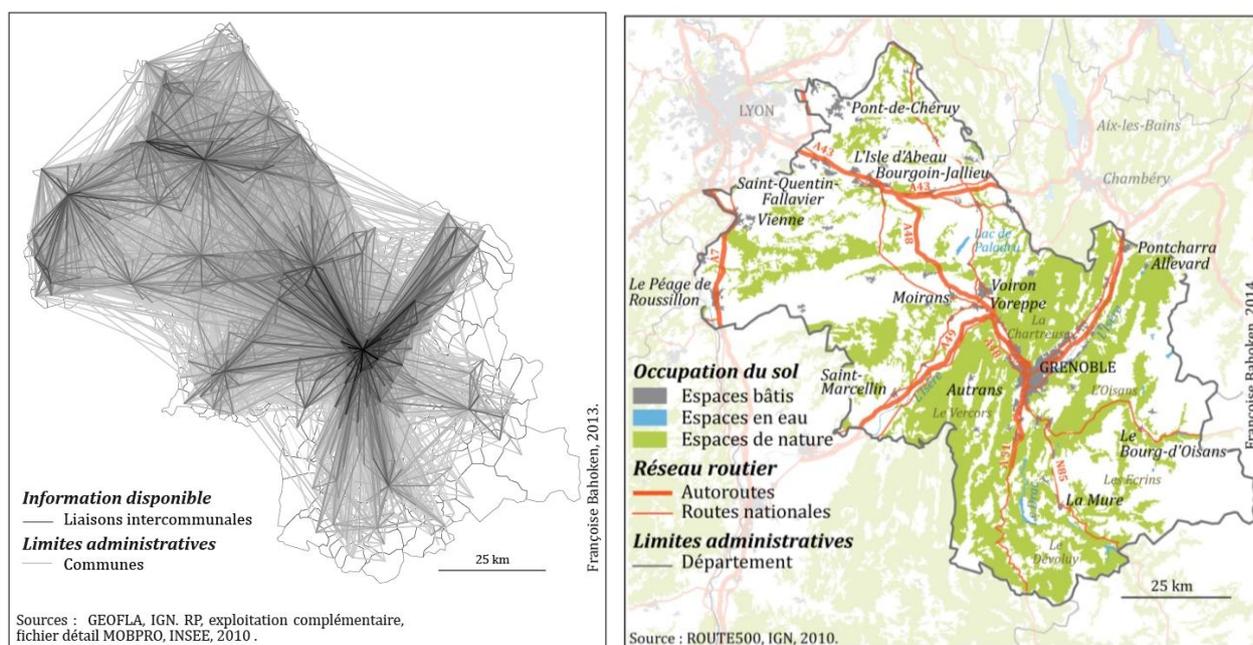
Les deux approches, sémiologiques et numériques, sont présentées successivement dans les deux sections suivantes.

7.1.1 L'approche sémiologique : le recours à la *Transparence*

Une manière simple et efficace de visualiser les différentes classes de distances parcourues par le flux, sans agir sur les valeurs de la matrice, consiste à mobiliser la *Transparence*.

La Figure 7-1 représente, à titre d'exemple, la cartographie des liaisons (Lij) intercommunales de l'Isère selon une *logique de flux* traditionnelle, en intégrant des variations de l'opacité de la teinte de manière inversement proportionnelle à la distance parcourue. Plus le figuré est court et plus opaque est sa teinte et inversement. N'ayant – volontairement – pas d'information sur la valeur de ces liens, il ne nous est pas possible de réaliser une légende qui renseigne sur le niveau de teinte associé à la longueur des liens. Pour comparer la vraisemblance du résultat obtenu, nous présentons sur la partie droite, de cette figure, une carte simplifiée de la topographie de l'Isère issue d'une sélection de données disponibles dans le fichier ROUTE500© de l'IGN (2010).

Figure 7-1. Cartographie des liaisons (Lij) de l'Isère



La variation de la valeur d'opacité de la teinte, en fonction de la *longueur* du lien, conduit à la perception de plusieurs systèmes de mobilités emboîtés les uns dans les autres. En effet, on distingue des liens courts dont l'opacité est la plus importante, et qui correspondent à des mobilités très locales, d'une vingtaine de kilomètres. Ces derniers caractérisent les flux localisés au cœur du Y grenoblois, coincé entre les massifs du Vercors à l'Ouest, de la Chartreuse au Nord, de Belledonne et de l'Oisans à l'Est. On devine aussi que ces mobilités professionnelles sont réalisées majoritairement par mode routier : les principaux axes de circulation qui y participent, très probablement, sont nettement visibles sur les cartes de cette Figure 7-1. Il s'agit de la rocade sud A480, qui dessine une première couronne fonctionnelle du sud-ouest au nord-est de la ville de Grenoble, mais aussi de l'autoroute A48 vers Voiron, puis vers les villes de Bourgoin-Jallieu et de l'Isle-d'Abeau situées plus au nord du département ; l'autoroute A41 vers la Savoie, en passant par Pontcharra et Allevard. Les pôles secondaires, typiques d'un second système de mobilités quotidiennes sont également perceptibles : il

s'agit, par exemple, de Saint-Marcellin et de Vienne, des communes situées respectivement à l'ouest et au nord-ouest du département, mais aussi du Bourg-d'Oisans au sud-est.

Afin d'analyser plus finement ces mobilités professionnelles grenobloises, il est possible de segmenter cette carte en autant de systèmes de mobilités (en raison de *l'effet de profondeur* des cartes de flux) dont le nombre dépendra de l'analyse. Il est également possible de raisonner selon une *logique de mouvement*.

En effet, nous avons vu dans le chapitre 6 que la réalisation d'une carte de flux implique de s'interroger sur la sémantique de l'information dont il est question (un flux spatialisé ou un mouvement) puisque leur construction cartographique est invariante. Dans ce contexte, en prenant en compte la distance parcourue, le questionnement concerne alors le fait de savoir s'il vaut mieux privilégier les flux qui sont en apparence faibles en valeur et en opacité, mais qui sont les plus nombreux en fréquence, ou bien les flux qui sont moins nombreux, mais dont la valeur et l'opacité sont les plus importantes. Ce questionnement nous renvoie à l'hypothèse n°3¹⁸⁹ du chapitre 6. Nous avons montré (section 6.3.2, page 275) qu'elle entraînait une surreprésentation abusive de la perception des petits flux, en raison, d'une part, du doublement de la distance ($S=F_{ij}.D_{ij}^2$) et d'autre part, d'une augmentation de *l'effet de couverture spatiale*. En effet, le fait que la cartographie directe du mouvement privilégie la perception des plus petits flux (les plus fréquents et les plus longs) au détriment des plus gros (les moins fréquents et les plus courts) peut être préjudiciable aux analyses réalisées sur des matrices de flux s'exprimant à courte distance. Pour éviter ce problème, il est possible d'améliorer la représentation correspondante du mouvement, dans le cadre de l'hypothèse n°3, en agissant au niveau sémiologique, comme nous l'avons montré dans le chapitre 5 (section 5.2, page 222).

L'objectif est alors d'agir sur la perception de la *surface* du figuré au niveau sémiologique, c'est-à-dire sur la quantité de teinte noire correspondant à la nuance de l'intensité (voir Figure 6-1, page 249) et non sur ses dimensions géographiques (*largeur, longueur, hauteur*). Dit autrement, la solution consiste à faire varier l'intensité perçue du mouvement pour compenser l'effet visuel de l'augmentation de la *largeur* des plus petits flux liée à l'application de l'hypothèse n°3, en mobilisant la *Transparence*. Il convient de noter que l'on aurait également pu mobiliser la *Valeur*, mais cela aurait conduit à sélectionner le nombre de figurés à représenter (une éventualité que nous verrons dans le chapitre suivant). Deux possibilités d'introduction de la *Transparence* sont envisageables, selon que la valeur d'opacité est proportionnelle ou est l'inverse de la distance ($1/D_{ij}$) :

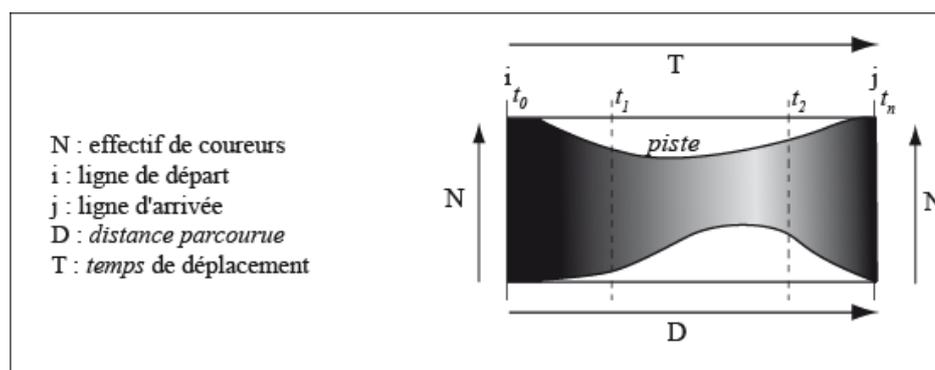
$(M_{ij})_A$: *largeur* = (M_{ij}) , *longueur*= D_{ij} et *Transparence*=1

$(M_{ij})_B$: *largeur* = (M_{ij}) , *longueur*= D_{ij} et *Transparence*= $1/D_{ij}$

Cet usage de la *Transparence* renvoie ainsi non pas à la perception d'une quantité de *mouvement* (M_{ij}), mais à une forme de densité qui pourrait correspondre à *l'intensité de mouvement spatial* (I_{ij}) du déplacement réalisé entre (i) et (j). Si nous reprenons l'analogie précédente des coureurs au départ d'une course de 100 m, cette mesure *d'intensité de mouvement* permet de prendre en compte le nombre de coureurs présent sur les lignes de départ et d'arrivée : plus ils sont nombreux et plus leur densité sur la ligne est élevée ou importante et inversement.

La Figure 7-2 illustre cette analogie du mouvement spatial avec une course à pied.

¹⁸⁹ Nous rappelons que la vraisemblance de l'hypothèse n°3 signifie que l'œil perçoit une surface géométrique $S=(F_{ij}.D_{ij}).D_{ij}$, d'où le doublement de la perception de la distance.

Figure 7-2. L'analogie avec la *densité de mouvement des coureurs*

D'après la Figure 7-2, on observe que l'hypothèse n°3 correspond en fait à **une probabilité de présence des coureurs sur la piste au cours du temps $[t_1-t_2]$: ces individus étant mobiles, ils occupent nécessairement une position dans l'espace (x, y) et dans le temps(t), pendant la durée d'observation $[t_0-t_n]$ de la course.**

De ce point de vue, si l'on représente cette mesure sur la bande qui représente la piste de course, cela revient bien à nuancer le degré d'intensité de la teinte de (M_{ij_1}) en fonction de la distance parcourue par les coureurs (D_{ij}) , qui correspond à (M_{ij_2}) , à rendre la teinte plus ou moins opaque. Etant donné que les coureurs sont en mouvement, l'intensité de noir sera appréciée lors du sifflet final (puisque l'on ne peut le saisir lorsqu'il se produit, la mesure étant rétrospective), au moment où le curseur s'arrête ; cette intensité de mouvement sera alors fonction de la distance parcourue. Dans notre cadre d'analyse, ce n'est pas tant la vitesse de déplacement entre les deux lignes (de départ et d'arrivée) ni la position des coureurs sur la piste qui importe (puisque l'information individuelle est perdue au profit de l'agrégat) que leur densité globale qui définit l'intensité de la teinte noire perçue.

La seconde possibilité de représentation de la distance parcourue passe par la réalisation d'une collection de cartes de flux segmentées par classes de distances parcourues.

7.1.2 L'approche cartographique: la collection de cartes, issues d'une segmentation de la matrice

L'intérêt de réaliser une collection de cartes tient dans la possibilité de représenter l'ensemble de l'information de flux, mais de manière segmentée sur plusieurs cartes. D'après J. Bertin qui s'exprimait au sujet des graphes, la collection de carte apporte une bonne réponse à l'impossibilité de produire une représentation graphique des flux sur une seule carte, et qui soit instantanément perceptible, lorsque plus de deux variables visuelles sont mobilisées : une planaire (x,y) et une rétinienne. Cette impossibilité pose d'ailleurs le problème de la représentation du mouvement de manière générale, de leur cartographie en particulier (Bertin, 1967). La complexité analytique de la figure implique que la perception de ses motifs demandera nécessairement un certain temps, lié à l'adaptation du regard. C'est pourquoi une manière d'y remédier consiste à segmenter l'information en k sous-parties.

L'idée de segmenter une information complexe en k sous-parties a été inventée par J.E. Marey (voir Encart 7-1), pour représenter le *mouvement* – une fonction par essence dynamique – sur un support passif ou statique.

Cette solution reprise par J. Bertin a donné lieu à la collection de cartes, aussi qualifiée de type bande-dessinée. L'auteur n'ayant toutefois pas précisé, à notre connaissance, de quelle manière cette segmentation devait être réalisée, nous faisons l'hypothèse que **les méthodes classiques de la**

discrétisation cartographique des séries de stocks continues sont applicables aux flux : les bornes des différentes classes pouvant justement être considérées comme les limites des cartes.

Dans les faits, lorsque les cartes sont présentées sous la forme d'une collection, c'est généralement pour donner à voir le changement ou la permanence du phénomène en fonction du temps, chacune des cartes correspondant alors à une date. Ici, la logique est différente puisqu'il s'agit de représenter la même série de données, observée à une date fixée, mais de manière segmentée.

Encart 7-1. Jules Marey et la segmentation graphique du mouvement

Le principe de la représentation graphique sous la forme d'une collection d'images a été mis au point par E.-J. Marey (1830-1904), médecin et physiologiste français dont une partie des travaux a porté sur la représentation graphique du « *mouvement à l'intérieur du corps* », puis des « *corps en mouvement* ». Sa thèse soutenue en 1857 sur la circulation sanguine est d'ailleurs considérée comme magistrale. Les travaux de l'auteur, dont une partie est disponible en ligne à la suite de l'exposition : *Le mouvement en lumière, Etienne-Jules Marey*, conçue et produite par la Maison du cinéma et la Cinémathèque française (<http://www.cinematheque.fr/marey/expo-ligne/> [lien vérifié en février 2015], constituent le fondement des méthodes graphiques actuelles de représentation des mouvements.

Poussé par l'idée de « *voir l'invisible* », en arrêtant ou accélérant le temps du mouvement, l'auteur s'attache à tous les décrire de manière graphique puis à l'aide d'outils, parmi lesquels la chronophotographie, un procédé inventé pour remédier aux limites de la représentation graphique plane que l'auteur considère comme « *l'écriture du mouvement* ». C'est à partir de la circulation du sang, puis de l'étude des muscles, que E.-J. Marey s'intéressera à la locomotion puis à l'observation du vol des insectes et des oiseaux, ou de la nage de poissons, sous une forme statique, dans un plan. L'intérêt de ses propositions porte, en premier lieu, sur la manière dont un graphique représentant le mouvement d'un mobile représenté par un point garde la mémoire de la trajectoire aux différentes étapes et ce, par la représentation de la trace du mobile ; en second lieu, sur le fait que la décomposition du mouvement par pas de temps consiste à superposer plusieurs images : une pour chaque période, l'ensemble correspondant à l'information totale. Outre ces propositions qui seront reprises ultérieurement, des diagrammes spécifiques pour la représentation de données temporelles ont été élaborés par l'auteur.

Source : Cinémathèque Française.

Cependant, plutôt que segmenter la matrice (Fij) pour obtenir une collection de cartes de flux, nous allons mobiliser une matrice des *distances géographiques* (Dij), qui sera discrétisée, de manière à ce que la collection de cartes corresponde à la représentation de différentes classes de distances parcourues par les flux.

7.1.2.1 Le principe de la segmentation des flux, en fonction des distances (Fij ← Dij)

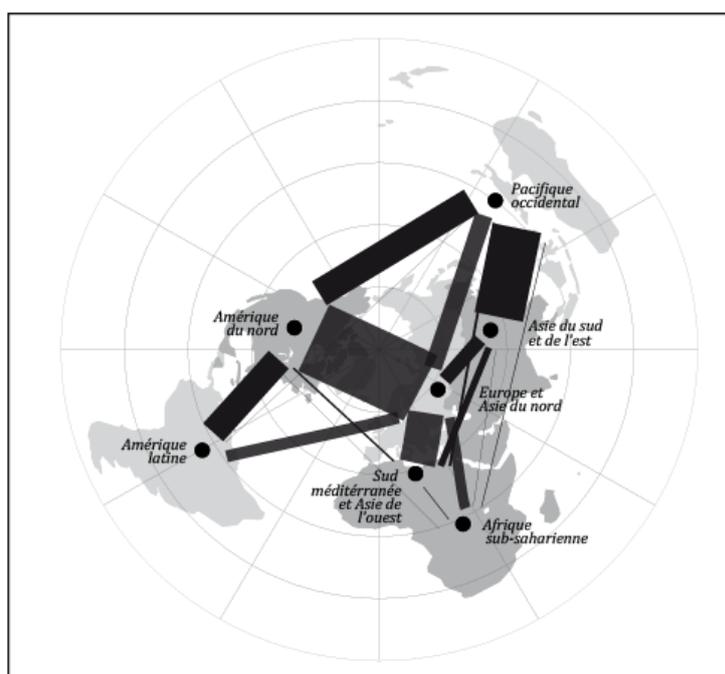
Nous mobilisons une matrice de distances géographiques de type (Dij), telle que définie dans la section 2 du chapitre 6 (page 258). Cette matrice est de type quantitatif continu : elle est définie sur un intervalle réel : $D_{ij} \in [(\text{Min}(D_{ij}) ; \text{Max}(D_{ij}))]$, en unités kilométriques. L'objectif est de la discrétiser en k classes de distances, dont les bornes serviront à segmenter la matrice (Fij), également en k classes. La difficulté réside donc dans l'estimation des valeurs des bornes (Min) et (Max) des k classes.

Deux approches sont envisageables. La première consiste à raisonner selon les dires d'experts de la thématique, tant pour ce qui est du choix de la métrique que de la valeur du ou des seuils, de manière à ce que la mesure d'éloignement retenue soit la plus pertinente possible. Pour les mobilités locales, il est, par exemple, possible de considérer le seuil de 100 kilomètres défini par J-P. Orfeuil, que nous

avons mentionné dans le chapitre 1. De la même façon, on sait que l'effet d'une gare ferroviaire sur les mobilités individuelles réalisées en transport en commun est maximal dans un rayon de 800 mètres autour de celle-ci. Une autre mesure possible de l'éloignement ferroviaire est celle du « *potentiel de contact* » entre deux lieux, défini par A. L'Hostis et al. (2004) comme le temps qui peut être passé à destination (j), compte tenu des temps de transport cumulés nécessaires pour faire l'aller-retour dans la journée entre (i) et (i). Les travaux réalisés en ce sens estiment ce temps à une durée de six heures de présence sur le lieu de destination, pour les villes les plus lointaines, avec un départ après 5h30 du matin et un retour, le soir, avant 23h30. Compte-tenu des difficultés liées à l'estimation d'un seuil unique de distance permettant de séparer les mobilités locales de mobilités à plus longue distance, indépendamment de leur type, nous faisons le choix d'appliquer une méthode de discrétisation statistique aux données de flux, de segmenter une matrice de flux (Fij), en (k) classes de distances (Dij) : $F_{ij} = D_{ij} \leftarrow F_{ij}$.

Pour illustrer notre propos, nous considérons une matrice¹⁹⁰ de flux mondiaux de passagers aériens entre sept grandes régions mondiales, dont la cartographie est présentée sur la Figure 7-3.

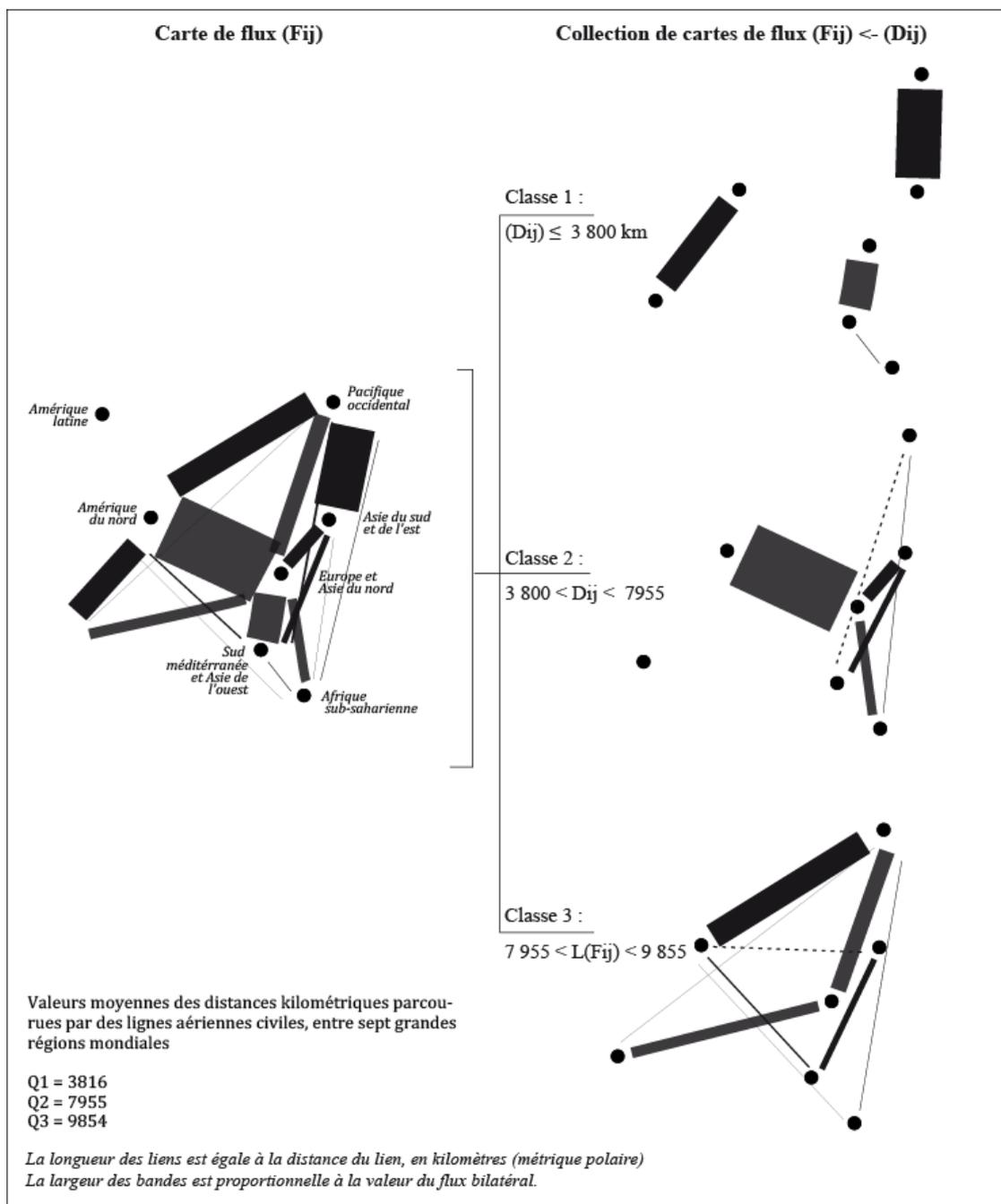
Figure 7-3. Présentation d'une matrice de flux mondiaux de passagers (OACI, 2000)



Le résultat cartographique issu de la segmentation de la matrice de flux correspondante est présenté sur la Figure 7-4 à laquelle on a ôté le fond de carte. La carte d'origine, qui couvre l'ensemble des flux, est présentée en trois classes de distances complémentaires (selon les quantiles). Elle correspond à la décomposition d'un phénomène complexe en un ensemble de cartes simples qui pourront être examinées soit simultanément, soit séparément. La question qui se pose alors est celle de la position des bornes de classes définissant l'univers de chaque carte ce qui passe, en cartographie statistique, par l'application d'une méthode de discrétisation.

¹⁹⁰ Nous reviendrons sur la construction de cette matrice dans la section 7.4.4.

Figure 7-4. Illustration de la segmentation d'une matrice de flux, pour une collection de cartes



Si le choix d'une métrique d'éloignement est important, le choix des seuils de distance permettant de réaliser une collection de cartes de flux peut, dans certains cas, s'affranchir de considérations empiriques et s'inscrire dans un raisonnement statistique plus général applicable à tous types de matrice de flux. La question du choix des bornes des classes de distance est, en effet, possible, en appliquant les méthodes de la cartographie statistique choroplèthe à celle de flux.

En cartographie statistique, la réalisation d'une collection de cartes issues d'une série statistique unique consiste à segmenter la série en un nombre k de plages de valeurs distinctes. La discrétisation cartographique correspondante sera appliquée sur la matrice des distances (et non sur celle des flux)

puisque l'objectif est de segmenter la matrice en fonction de la distance parcourue. L'intérêt du recours à une discrétisation cartographique tient surtout à la possibilité de comparer des cartes entre elles, en utilisant des paramètres statistiques. De nombreuses possibilités sont envisageables et C. Cauvin et al. (1978) en ont proposé des typologies qui sont toutes fondées sur des données de stocks, dans le cadre de la cartographie choroplèthe. Il n'existe pas, à notre connaissance, d'examen spécifique des méthodes de discrétisation adapté aux objets linéaires et donc aux flux.

Nous privilégions deux grandes familles de méthodes de discrétisation : les moyennes emboîtées et les quantiles. L'intérêt est de ne pas postuler sur la normalité de la série, tout en restant robuste. La standardisation, fondée sur l'utilisation de la moyenne et de l'écart-type, pourtant très utilisée pour les données de stocks, a été écartée, car elle s'est avérée inadaptée au cas des flux. En effet, la transformation des valeurs (de flux qui sont fortement dissymétriques) conduit à un ajustement imparfait des données et se révèle surtout utile pour les séries gaussiennes, comme nous le verrons dans le chapitre 8.

Les deux méthodes de discrétisation sélectionnées sont présentées successivement dans les deux sous-sections suivantes. Leur choix tient au fait qu'elles peuvent être mobilisées pour réaliser une collection de cartes visant à être comparées entre elles. Si les moyennes emboîtées permettent de mesurer l'écart des valeurs de flux à la moyenne de chaque classe, les quantiles comparent l'ordre pris par les unités spatiales, leur position ou leur rang.

7.1.2.2 Les classes fondées sur les moyennes emboîtées

La méthode des moyennes emboîtées s'appuie sur la valeur de moyenne arithmétique (\bar{x}) pour discrétiser la série en un nombre pair de classes.

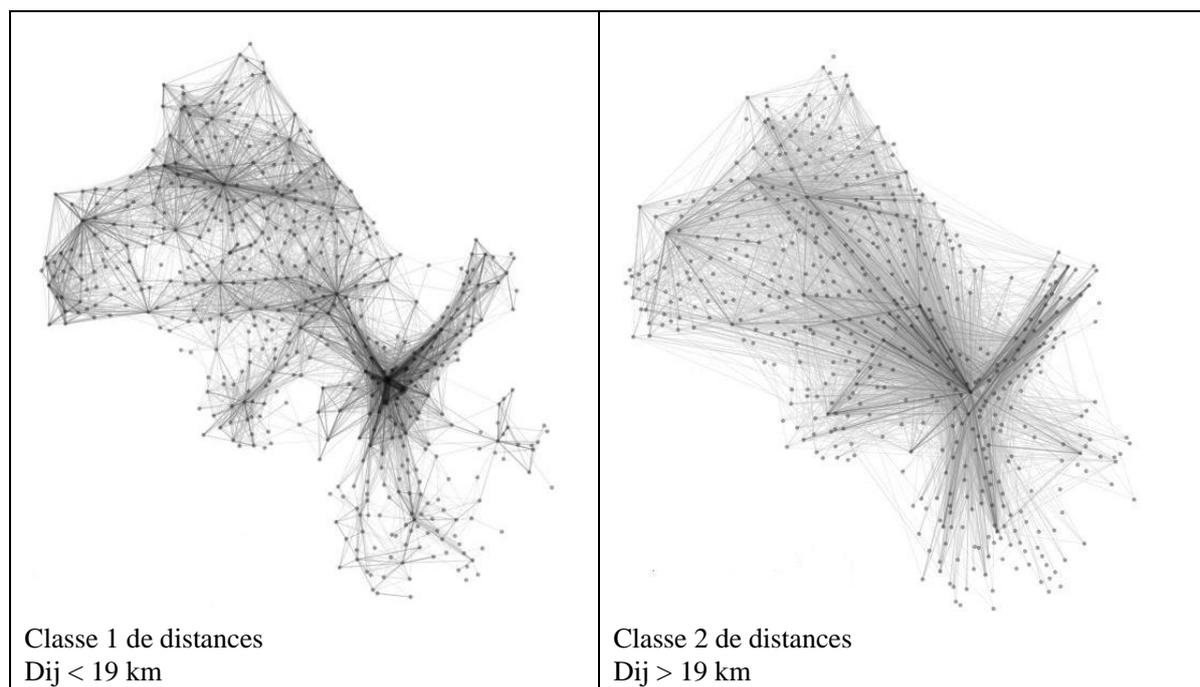
Le Tableau 7-1 illustre la construction de deux et quatre classes de distances parcourues par les flux de l'Isère, segmentées selon les moyennes emboîtées.

Tableau 7-1. Construction des bornes de classes de distances (Dij) de l'Isère, selon les moyennes emboîtées

Classe	Dichotomie	Discrétisation en 4 classes
Classe 1	[0,9 ; 19 [[0,9 ; 10 [
Classe 2	[19 ; 104 [[10 ; 19 [
Classe 3		[19 ; 32 [
Classe 4		[32 ; 104 [

Les deux figures suivantes présentent deux collections de cartes, respectivement de deux et quatre images, issues du Tableau 7-1. On observe que les motifs des flux varient en fonction de la classe de distance parcourue. La dichotomie, qui segmente la matrice en deux parties, autour de la moyenne des distances, semble bien discrétiser les flux locaux, comme illustré sur la figure suivante.

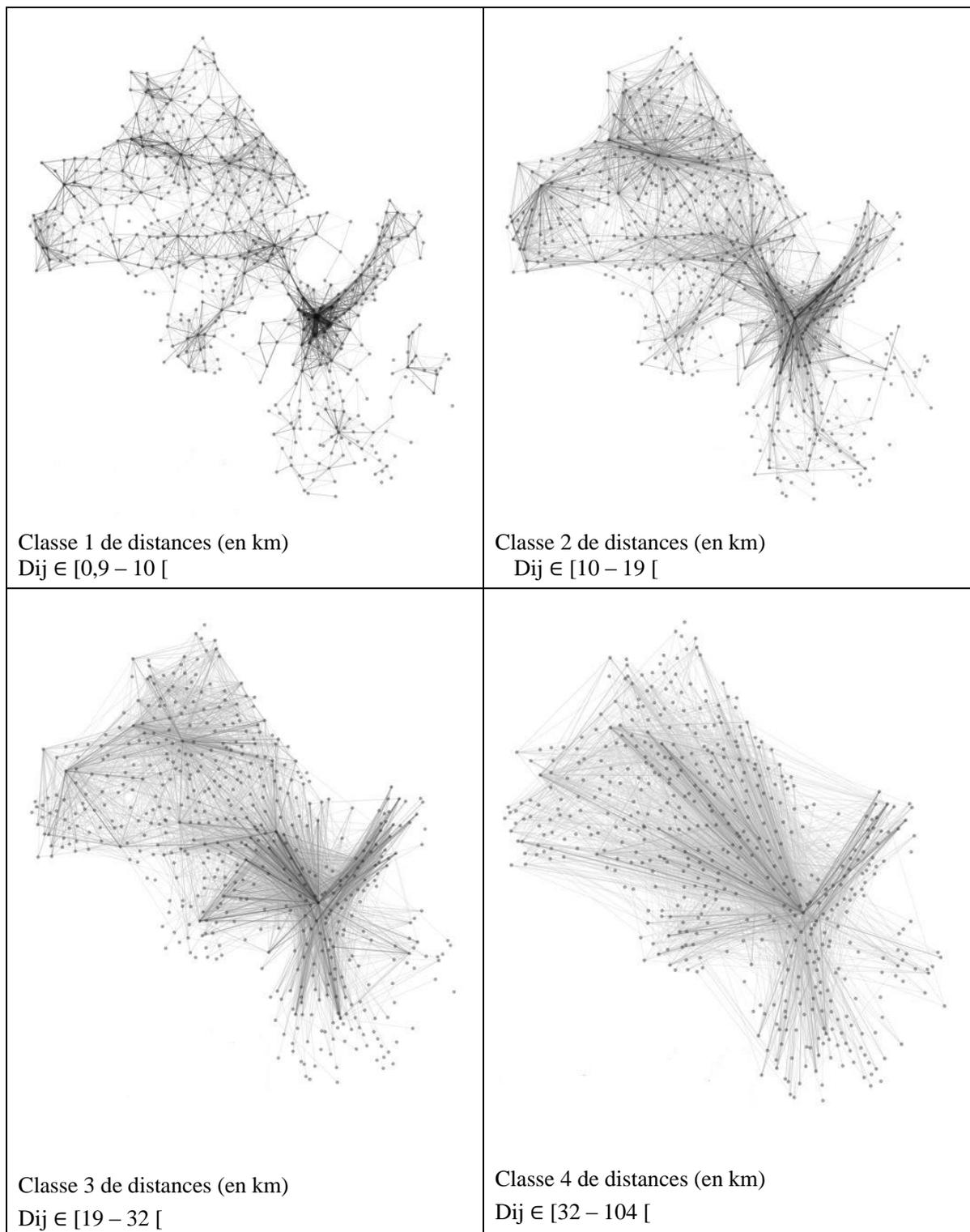
Figure 7-5. Collection de cartes du volume bilatéral de l'Isère (Fij⁺) issue d'une discrétisation de la matrice des distances (Dij) selon les moyennes emboîtées



Source : RP, exploitation complémentaire, fichier de détail MOBPRO, INSEE, 2010.

L'ensemble de quatre cartes suivantes affine ce résultat, en mettant en évidence la relation inverse entre la valeur du flux et la distance parcourue, autrement dit l'ajustement négatif des valeurs de flux à mesure que la distance parcourue augmente.

Figure 7-6. Collection de cartes du volume bilatéral de l'Isère (Fij⁺) issue d'une discrétisation de la matrice des distances (Dij) selon les moyennes emboîtées



Source : RP, exploitation complémentaire, fichier de détail MOBPRO, INSEE, 2010.

7.1.2.3 Les classes fondées sur les quantiles

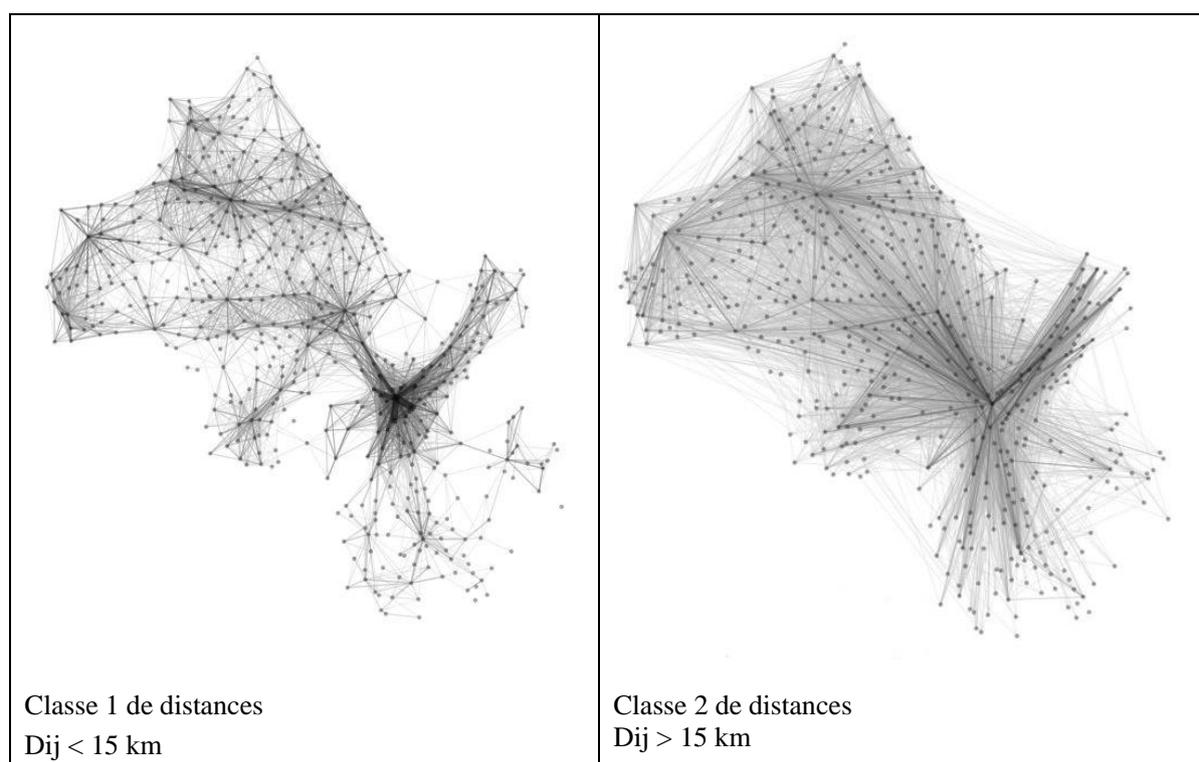
La seconde possibilité de discrétisation statistique des valeurs de distance parcourue par le flux, consiste à raisonner selon les quantiles. L'intérêt de cette méthode est de ne pas postuler la normalité de la série, elle peut donc être mobilisée directement pour discrétiser les distances.

Le recours aux quantiles permet de comparer la position des valeurs des couples de lieux par rapport à la médiane et non plus la moyenne des distances. Elle segmente la série en classes d'effectifs égaux et permet d'analyser leur ordre, en fonction de classes fondées sur les valeurs de quartiles des distances géographiques parcourues. Ces classes, toujours multiples de deux, sont définies de la manière suivante :

- Classe n°1 : $[\text{Min}(F_{ij}) - Q_1(D_{ij} \leftrightarrow F_{ij}) [$
- Classe n°2 : $[Q_1(D_{ij} \leftrightarrow F_{ij}) - Q_2(D_{ij} \leftrightarrow F_{ij}) [$
- Classe n°3 : $[Q_2(D_{ij} \leftrightarrow F_{ij}) - Q_3(D_{ij} \leftrightarrow F_{ij}) [$
- Classe n°4 : $[Q_3(D_{ij} \leftrightarrow F_{ij}) - \text{Max}(F_{ij})]$

La Figure 7-7 présente une collection de cartes du volume bilatéral des navettes de l'Isère (F_{ij}^+) issue de la segmentation de la matrice des distances (D_{ij}), en utilisant les valeurs de quantiles suivantes : $Q_1=8,4$ km, $Q_2=15,0$ km et $Q_3=25,2$ km. Comme précédemment, la distance parcourue minimale est de 0,9 km (arrondie), la distance maximale de 103,9 km et la médiane 15,0 km.

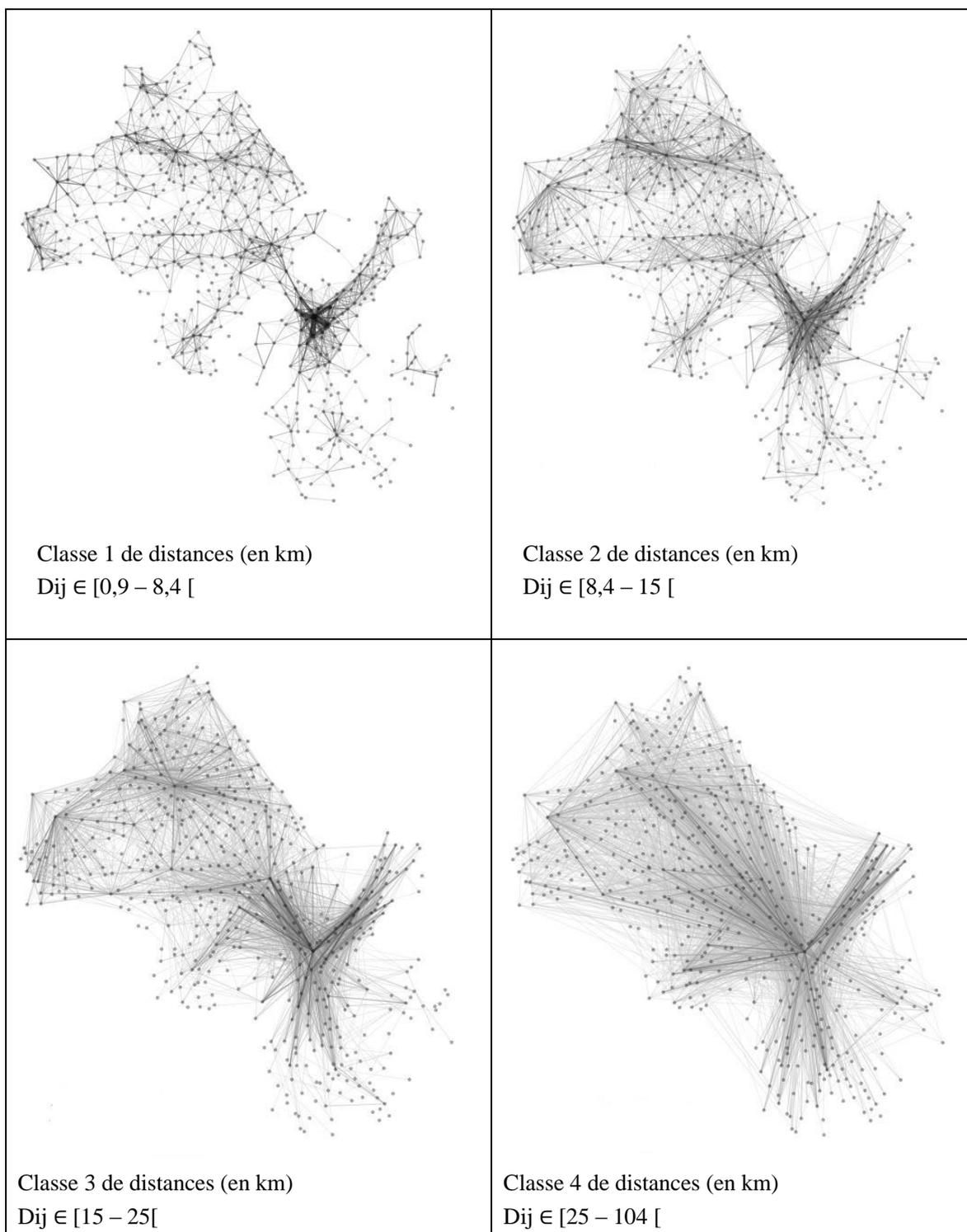
Figure 7-7. Collection de deux cartes du volume bilatéral des flux (F_{ij}^+) de l'Isère, issue d'une segmentation de la matrice des distances (D_{ij}) en fonction de la médiane



Source : RP, exploitation complémentaire, fichier de détail MOBPRO, INSEE, 2010.

Comme précédemment, les cartes demeurent très complémentaires, elles sont dans l'esprit de la collection de cartes.

Figure 7-8. Collection de quatre cartes du volume bilatéral des flux (F_{ij}^+) de l'Isère, issue d'une segmentation de la matrice des distances (D_{ij}) selon les quartiles



Source : RP, exploitation complémentaire, fichier de détail MOBPRO, INSEE, 2010.

Le fait que la méthode génère des classes d'effectifs égaux peut poser différents problèmes qui résultent de la discrétisation d'une série (F_{ij}), dont la segmentation est construite sur une série exogène (D_{ij}). En effet, l'une des conséquences de la réalisation indirecte d'une discrétisation réside dans la forte variation de la fréquence des valeurs de chacune des classes, en raison de la relation inverse entre

valeur de flux et distance. C'est la raison pour laquelle, les cartes lisibles sur la Figure 7-8 sont celles dont la distance est inférieure au troisième quartile ($D_{ij} < 15\text{km}$).

Sur le plan thématique, la réalisation de telles collections de cartes fondées sur les quartiles permet de distinguer les flux courts, qui parcourent de faibles distances, des flux longs : ceux qui parcourent de longues distances. La polarisation de ces mobilités quotidiennes de l'Isère autour des principaux pôles d'emplois est d'ailleurs visible sur l'ensemble des cartes. Cependant, si la première image (des flux qui parcourent les classes de distances les plus faibles, inférieures à 8,4 km) décrit l'aire d'influence du pôle grenoblois et des pôles d'emploi secondaires, les cartes suivantes mettent l'accent sur les flux observés entre les principaux pôles d'emploi qui font fi de la topographie contrainte des lieux.

Les volumes bilatéraux de flux les plus importants concernent ainsi des actifs qui vont travailler dans les communes situées au cœur de l'agglomération, sur le replat de la cuvette grenobloise ou dans l'aire d'influence de principaux pôles d'emplois. Nous verrons dans la section 7.3 qu'ils correspondent au zonage territorial particulier des bassins de vie, qui sont polarisés.

La proximité entre les lieux est généralement introduite dans l'analyse cartographique des flux sous une forme continue, par le recours à la distance géographique parcourue (D_{ij}). Or, en fonction de la thématique, il peut être intéressant de mobiliser d'autres distances empiriques traduisant des formes de proximités différentes de (D_{ij}), qui varieront aussi en fonction du type de mesure. A la métrique continue peut, effectivement, être substituée une mesure discrète : binaire ou ordinale. Ces mesures caractérisent une acception de l'espace fondée sur la notion de proximité. Nous interprétons cette proximité (ou l'éloignement) d'un point de vue géométrique, par la notion de contiguïté (C_{ij}).

7.2 La métrique de proximité : le *voisinage* spatial binaire (C_{ij}) ou ordinal (CK_{ij})

L'introduction d'un critère de distance appréhendé par la proximité (ou l'éloignement) consiste à raisonner du point de vue du voisinage topologique des lieux, la proximité est alors appréhendée sous une forme discrète et non plus continue, comme dans le cas précédent.

Cette logique conduit à intégrer un effet du problème des unités spatiales modifiables (MAUP) lié au fait que les processus géographiques sont dépendants de l'échelle d'observation et de la résolution du maillage, autrement dit des *limites de zones* partitionnant l'espace d'étude à franchir par le flux. Cet effet lié au zonage peut être interprété selon deux approches. La première et la plus répandue appréhende la limite comme une discontinuité territoriale, ou bien en tant que variable d'ajustement des flux que l'on estime par la modélisation : la *limite de zone* aura alors un rôle positif, dans le cas des effets dits de relation préférentielle, qui facilitent les échanges ; ou négative, un rôle de frein dans les effets de barrière déjà mentionnés. Une seconde approche consiste à interpréter la présence d'une *limite de zone* comme une mesure discrète et binaire, dans un premier temps, de la proximité immédiate de deux unités spatiales. Deux mailles seront alors considérées comme proches si elles partagent une frontière commune, c'est-à-dire si elles sont contigües ou adjacentes.

La notion de contiguïté des unités spatiales permet de considérer le voisinage d'un lieu dont l'intérêt est d'autant plus important en analyse spatiale qu'il engendre (ou non) un effet d'interaction spatiale. Ce principe correspond aussi à l'une des hypothèses des modèles gravitaires selon laquelle **l'intensité des flux est d'autant plus importante que les lieux sont proches et, dans le cas présent, qu'ils sont contigus et partagent une frontière commune ($C_{ij}=1$)**. « *La contiguïté est la propriété d'objets ou de lieux qui sont voisins, qui se touchent, parce qu'ils ont une frontière commune s'il s'agit de mailles ou de zones, ou qui sont connectés par une ligne, une voie de communication s'il s'agit de nœuds dans un réseau* » (Pumain, Saint-Julien, 2005 : 29).

C'est pourquoi W. Tobler (1987) suggère le passage à la contiguïté ($C_{ij}=1$) de sorte que les migrants soient « [...] réorientés pour ne passer que dans des zones adjacentes. Ceci peut être réalisé de

différentes manières, mais le résultat final est le même ; la carte de mouvement ne contient pas de bandes ou de flèches qui se chevauchent » (Tobler, 1987 : 162)

Dans l'exemple des communes de l'Isère, nous considérons le voisinage d'un lieu en termes de frontière commune. Sur le plan cartographique, mobiliser le voisinage pour cartographier des flux conduit à mettre en évidence un phénomène d'osmose (Bertin, 1973 : 351), caractéristique des flux qui s'expriment entre des zones partageant une frontière commune.

De ce point de vue, la matrice de voisinage que nous mobiliserons sera calculée sur un maillage territorial surfacique et non ponctuel. Nous allons interpréter la mesure du voisinage (immédiat ou non) comme le caractère de ce qui est limitrophe : deux communes sont voisines si elles sont séparées par une frontière de longueur non nulle ($C_{ij}=1$). Cette mesure sera ensuite généralisée à l'existence d'au moins une de ces frontières, archivées dans la matrice $(CK_{ij})=\{1,2,3,4,n\}$, permettant de tenir compte d'une proximité dite de contact ($CK_{ij}=1$) ou de l'éloignement en nombre minimum de frontières à franchir. Ces deux possibilités sont successivement examinées dans les sous-sections suivantes. Nous présentons dans l'Annexe A.2 associée à ce chapitre 7, une analyse cartographique de la proximité des zones, selon la métrique du *plus proche voisin* (elle fait également la section 3 de l'annexe R-3).

7.2.1 La contiguïté simple ($C_{ij}=1$)

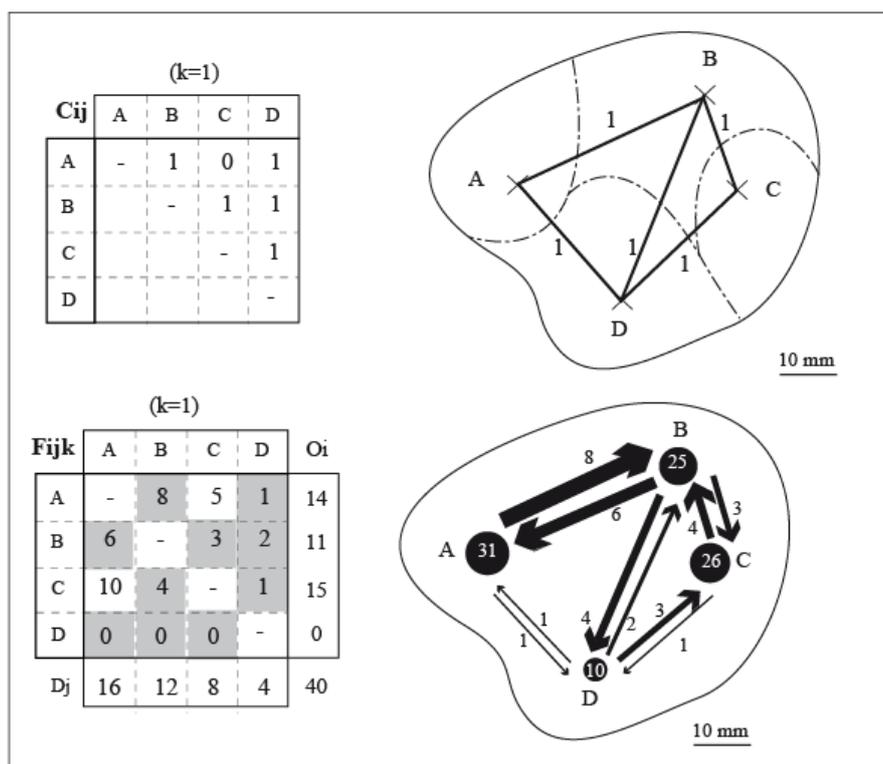
La mesure d'éloignement binaire correspond au cas où la matrice qui traduit l'espace des lieux est fondée sur la notion de contiguïté (C_{ij}), c'est-à-dire sur une notion de voisinage spatial tenant compte des frontières. Elle décrit la proximité ou l'éloignement des lieux en fonction de la position des *limites de zone*, plus précisément de l'existence d'une seule frontière commune entre deux unités spatiales d'origine et de destination. Elle est de fait composée de deux modalités qui renseignent sur l'existence de cette frontière commune ($C_{ij}=1$ ou non ($C_{ij}=0$). On parle alors généralement de degré de contiguïté d'ordre 1.

Contrairement à la mesure continue de l'éloignement (D_{ij}), il n'est pas nécessaire de rechercher d'éventuels seuils d'éloignement (ou de proximité) à représenter puisque la cartographie de la matrice de flux (F_{ij}) réduite au voisinage ($F_{ij} \leftarrow (C_{ij})$) se décomposera alors naturellement en deux cartes : la première décrira les flux réalisés entre des lieux proches ou voisins ($C_{ij}=1$) et inversement pour la seconde, qui concernera les flux qui se produisent entre des lieux éloignés ou non voisins ($C_{ij}=0$). Il sera dès lors possible de cartographier des migrations intercommunales (ou internationales), en utilisant la variable contiguïté (C_{ij}) qui renseigne sur la présence d'une frontière de longueur non nulle à franchir entre deux communes (ou deux pays).

C'est W. Tobler (1987 b) qui, le premier, suggéra l'utilisation d'un critère simple de contiguïté des zones égal à ($C_{ij}=1$), afin de ne représenter que les flux qui se produisent entre des zones voisines, considérant qu'ils correspondent à des mobilités locales. La sélection des flux en fonction de ce critère présente l'avantage d'apporter une information que l'on ne possède théoriquement que partiellement. En effet, les mobilités intra zonales sont en pratique perdues lorsque la diagonale de la matrice est portée à zéro. La cartographie des seuls flux qui se produisent entre des zones contigües conduit à un changement d'échelle qui permet de retrouver, certes partiellement, une information sur les flux intra zonaux, tout en conservant l'information sur les flux inter zonaux.

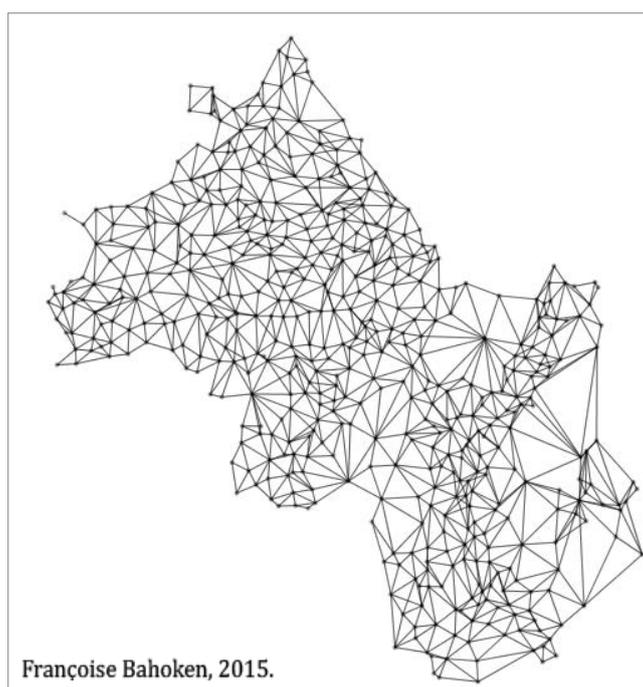
La Figure 7-9 illustre la cartographie de flux entre des régions limitrophes sur notre matrice exemple, elle décrit sur sa partie haute, le graphe de l'ensemble des relations répondant au critère (ou graphe de voisinage spatial d'ordre 1) qui correspond à la première étape de la procédure générale (C_{ij}). La seconde étape consiste à appliquer une variable de restriction à (F_{ij}), correspondant à (C_{ij}), afin de ne représenter que les flux qui sont réalisés entre des lieux limitrophes (partie basse de la Figure 7-9).

Figure 7-9. Illustration de la cartographie de flux entre des zones limitrophes ($C_{ij}=1$)



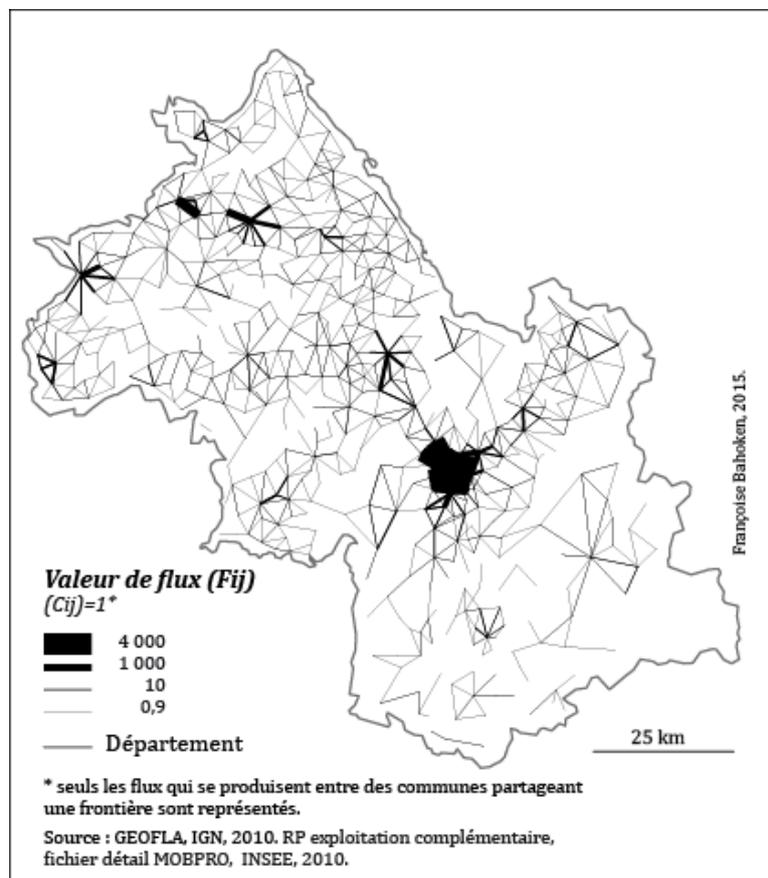
La Figure 7-10 présente le graphe de voisinage ($C_{ij}=1$) des communes de l'Isère. Elle décrit un graphe planaire valué de lecture aisée (sans *effet-spaghetti*), dont l'utilisation pour cartographier des flux intercommunaux entraînera nécessairement une réduction drastique de l'information, comme illustré sur la Figure 7-11.

Figure 7-10. Graphe de voisinage ($C_{ij}=1$) des communes de l'Isère



La Figure 7-11 présente la cartographie de la matrice des flux (F_{ij}^+) de l'Isère réduite aux seuls flux réalisés entre des communes proches, adjacentes.

Figure 7-11. Cartographie des flux (F_{ij}^+) \leftarrow ($C_{ij}=1$) entre les communes limitrophes de l'Isère



La cartographie de la matrice complémentaire à celle de la Figure 7-11 – qui concerne de fait les flux qui se sont produits entre des unités spatiales non limitrophes – continuera à subir *l'effet spaghetti*. Pour la représenter, il est possible de généraliser la mesure binaire du voisinage en la transformant en valeur ordinale, où $k=1, 2, 3, \dots, n$ formant ainsi (CK_{ij}). L'intérêt du recours à (CK_{ij}) est de permettre une cartographie des flux locaux, présentant une portée supérieure à l'ordre 1, sans toutefois représenter les flux à longue distance.

Nous proposons d'examiner la cartographie de [$F_{ij} \leftarrow (CK_{ji})$] dans la sous-section suivante en présentant le fonctionnement de la méthode de sélection sur notre matrice d'exemple local que nous allons étendre à huit lieux, avant de l'appliquer aux flux de l'Isère (F_{ij}).

7.2.2 La contiguïté ordinale (CK_{ij})

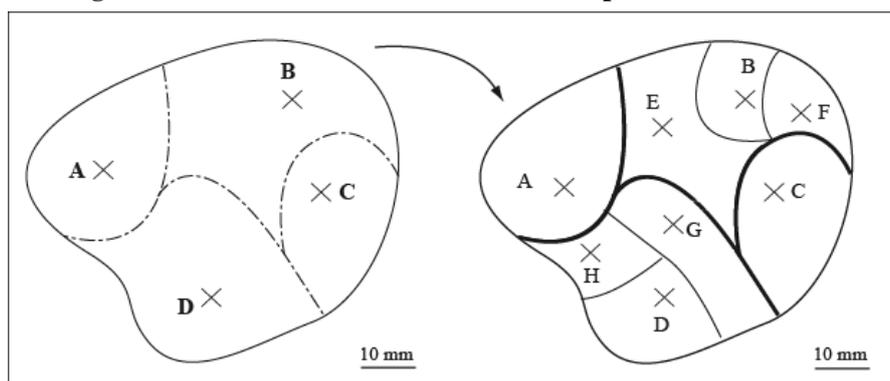
L'extension de la portée du voisinage (C_{ij}) de ($k=1$) à ($k=n$) a notamment été introduite au cours des années 1960 par L. Lebart (1969), pour analyser l'auto corrélation spatiale. Sur le plan formel, cette mesure permet de considérer des proximités d'ordre ($k=2, 3, \dots, n$) qui sont issues du calcul de plus courts chemins. La suggestion de son application à la cartographie des flux a été effectuée pour la première fois, à notre connaissance, par W. Tobler (1987b), à la suite de sa proposition précédente, pour un nombre de limites de zones à franchir par le flux d'une valeur ($k=4$). Reprise ensuite par D. Dorling (1991), elle fait l'objet d'une vive critique par ce dernier qui considéra la procédure inopérante suite à son application à des mobilités britanniques « *Une autre solution qui a été suggérée*

par Tobler est de réorienter les flux en fonction du plus court chemin à travers des zones contigües. Cela a été testé, mais il a été trouvé pour la Grande-Bretagne que l'effet était de produire une image inverse de la réalité. De nombreux flux sont émis du centre du pays et les relations entre les métropoles et leurs arrière-pays ont été inversés. L'idée a été abandonnée ».

Dans le cadre d'analyse des flux, nous pensons que cette proposition correspond à une variante¹⁹¹ intéressante de la cartographie des flux en classes de distances vues précédemment. Cela permet effectivement d'intégrer le fait que « [...] les interactions spatiales peuvent se propager dans l'espace de proche en proche, et pour l'analyse des ressemblances ou de la diffusion spatiales, on utilise souvent des matrices de contigüité d'ordre deux, ou d'ordre trois, ou d'ordre n quelconque, qui représentent les possibilités de contact ou d'interaction entre des nœuds dans un réseau ou entre les mailles d'une grille, lorsqu'on accepte de franchir des écarts plus grands [...]. On peut utiliser des matrices de contigüité pour calculer des distances topologiques, dites « mesures d'écarts, entre deux mailles quelconques d'une grille ou entre deux nœuds d'un réseau. On mesure alors le nombre de frontières à traverser, ou de segments à parcourir, pour se rendre d'un lieu à un autre. » (Pumain, Saint-Julien, 2005 : 31). La mesure de (CK_{ij}) étant ordinale, elle prend des valeurs comprises entre 1 et $+\infty$, si le graphe est non connexe.

Pour illustrer sa construction, nous mobilisons notre matrice d'exemple local, que nous étendons à huit lieux, afin de pouvoir disposer de plus de possibilités de sélection (voir Figure 7-12).

Figure 7-12. Présentation de la matrice exemple local à huit lieux



La construction de la matrice (CK_{ij}) correspondante est présentée sur la Figure 7-13.

Figure 7-13. Illustration de la construction d'une matrice de contigüité ordinale (CK_{ij})

		avec $k=1$										avec $k=1...3$							
CK _{ij}		A	B	C	D	E	F	G	H	CK _{ij}		A	B	C	D	E	F	G	H
A	-	0	0	0	1	0	1	1		A	-	2	2	1	3	1	1		
B	0	-	1	0	1	1	0	0		B	2	-	1	3	1	1	2	3	
C	0	1	-	0	1	1	1	0		C	2	1	-	2	1	1	1	2	
D	0	0	0	-	0	0	1	1		D	2	3	2	-	2	3	1	1	
E	1	1	1	0	-	0	1	0		E	1	1	1		-	2	1	2	
F	0	1	1	0	0	-	0	0		F	3	1	1	3	2	-	2	3	
G	1	0	1	1	1	0	-	1		G	1	2	1	1	1		-	1	
H	1	0	0	1	0	0	1	-		H	1	3	2	1	2	3	1	-	

¹⁹¹ Il existe d'autres variantes de la contigüité (Tour / Reine), mais elles sont surtout utiles sur des maillages réguliers ou de type grille (voir l'Annexe associée à ce chapitre 7).

Cette Figure 7-13 distingue le cas de $(CK_{ij}=1)$ – qui est à comparer avec $(C_{ij}=1)$ – et celui de $(CK_{ij}=n)$, placées respectivement sur les parties gauche et droite de la figure.

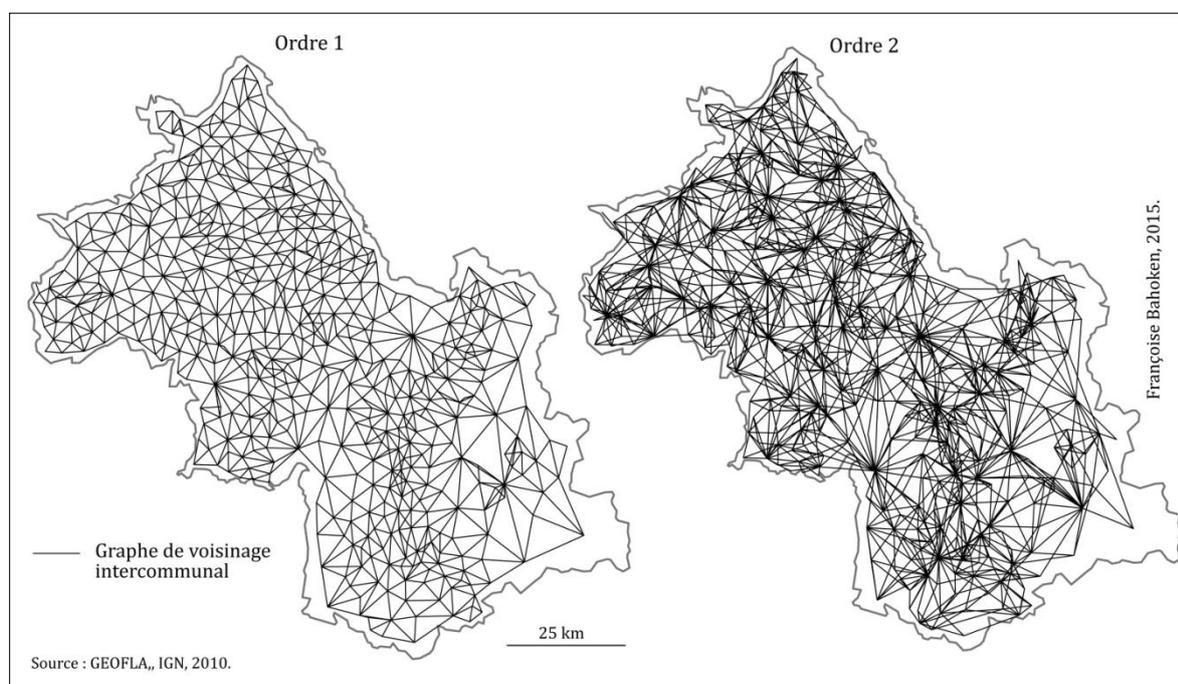
On observe que la matrice $(CK_{ij}=1)$ est égale à $(C_{ij}=1)$ ¹⁹², les deux matrices étant construites en adoptant le même raisonnement. (CK_{ij}) compte le même nombre de lignes et de colonnes, elle est nécessairement symétrique – tout comme C_{ij} – et sa diagonale principale est vide¹⁹³. La différence entre les deux matrices de la Figure 7-13 tient au rang d'éloignement considérée (1, 2, ..., n).

Sur le plan méthodologique, une manière de calculer les valeurs de (CK_{ij}) consiste à faire l'hypothèse que **l'existence d'un lien entre deux lieux marque la présence d'une frontière, ce qui permet de mobiliser la mesure de la distance géodésique définie sur le graphe de voisinage**. Cette *distance géodésique* s'exprime par un nombre de liens franchissant une limite de zone qu'il convient de parcourir pour aller d'un lieu (i) d'origine à un lieu (j) de destination.

La cartographie de la matrice de flux ainsi réduite au rang d'éloignement géodésique considéré $(F_{ij}) \leftarrow (CK_{ij})$ pourra alors se faire successivement soit entre les communes voisines à l'ordre 1, ce qui revient à $(C_{ij}=1)$, soit à l'ordre k $(CK_{ij}=k)$.

La Figure 7-14 présente les graphes de voisinage (CK_{ij}) des communes de l'Isère pour $(k=1)$ et $(k=2)$.

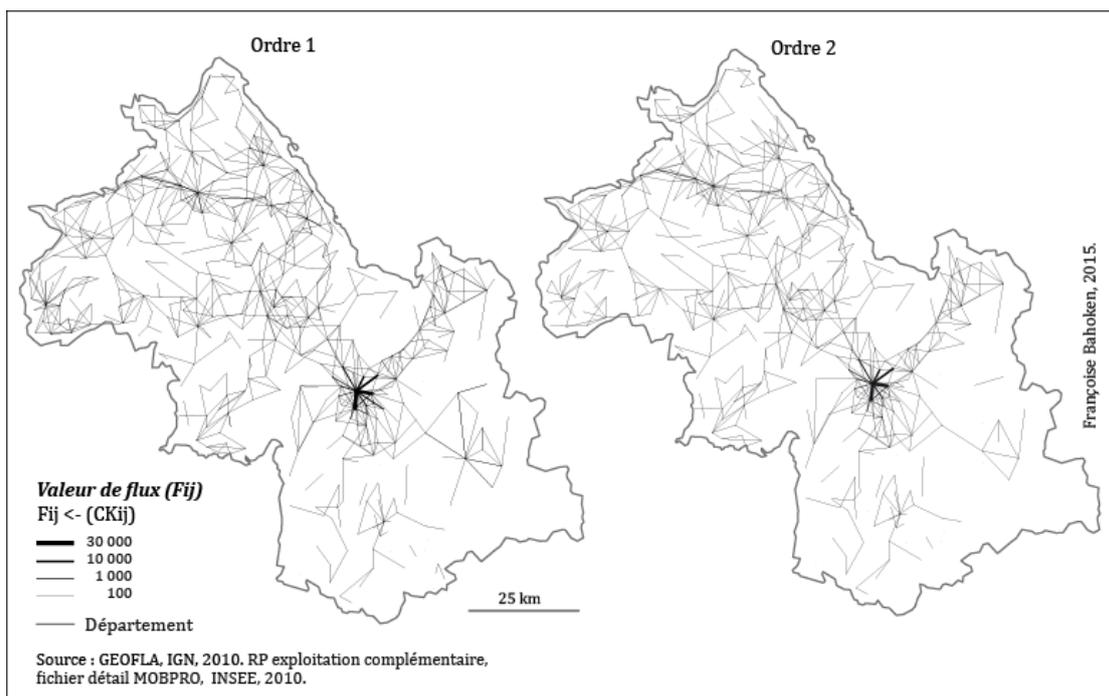
Figure 7-14. Graphes de voisinage des communes de l'Isère (CK_{ij}) avec $k=1$ et $k=2$



Les cartes de flux réduites correspondantes à ces deux graphes sont présentées sur la figure suivante.

¹⁹² Pour s'en rendre compte, il convient de comparer terme à terme les matrices de la Figure 7-13 où les cases présentant la valeur 1 sont identiques.

¹⁹³ Dans une matrice de contiguïté classique, la diagonale principale est renseignée par le degré des lieux (une mesure de l'analyse des réseaux correspondant au nombre de liens entrants ou sortants).

Figure 7-15. Cartographie des flux de l'Isère ($F_{ij} \leftarrow [(CK_{ij}), k=1 \text{ et } k=2]$)

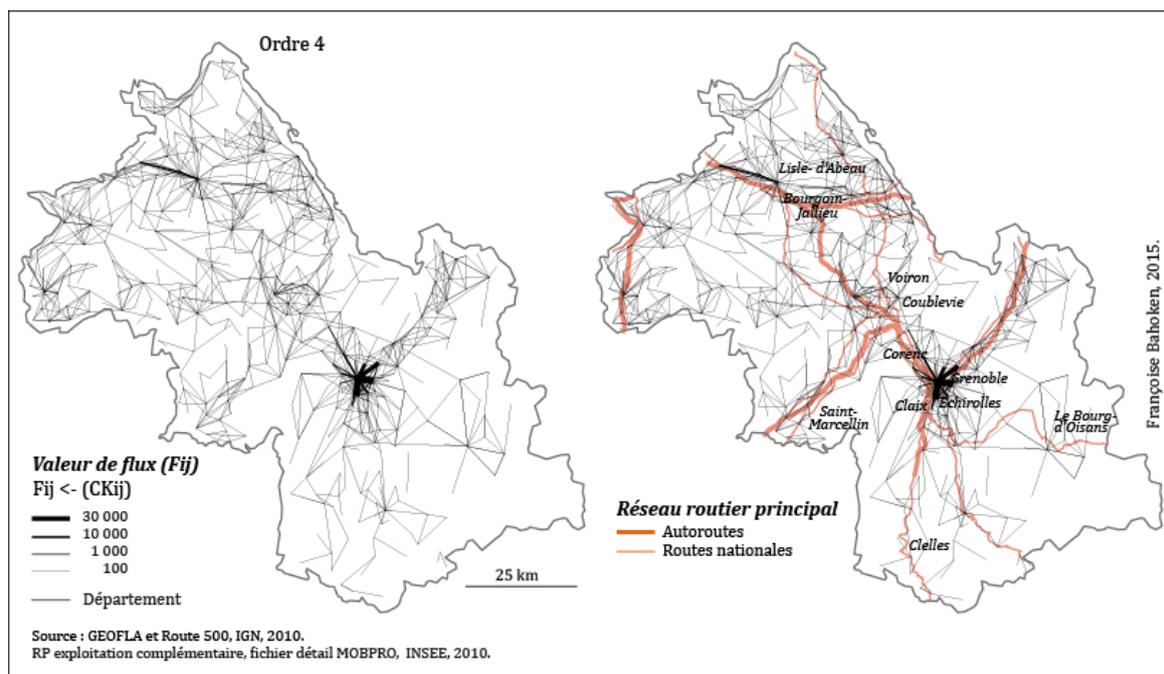
Comme dans la section 7.1, il sera possible de réaliser une collection de cartes des différents ordres de voisinage ou de n'en présenter qu'une seule, par exemple l'ordre 2 : $(CK_{ij}, k=2)$. Cela revient à cartographier les flux qui s'expriment à une portée de 1 et de 2, dont la représentation est plus intéressante d'un point de vue cartographique : elle permet de mettre en évidence l'augmentation de l'éloignement qui se traduit par l'accroissement de la densité de la matrice.

Si les images des graphes de voisinage (voir Figure 7-14) sont bien distincts, celles de flux présentées ci-dessus sont assez similaires, car les valeurs de portée de la distance d'ordre 1 et 2 diffèrent peu, les unités spatiales de l'Isère étant par endroits de très petite taille. Pour percevoir plus nettement l'effet de la réduction de la matrice de flux en fonction d'un critère de voisinage, il convient de choisir une portée plus importante, l'ordre 4, par exemple, d'ailleurs suggéré par W. Tobler. La carte de flux de l'Isère correspondante est présentée sur la Figure 7-16.

La Figure 7-16 (à gauche) décrit les motifs de flux de l'Isère d'une portée de 4 limites de zones $(CK_{ij}, k=4)$, celle de droite replace ces motifs au regard du réseau routier principal. L'utilisation d'un critère de contiguïté ordinal met bien en évidence l'expansion des flux de proche en proche, comme l'a suggérée par W. Tobler (1987b), en suivant le maillage routier principal dont on perçoit convenablement l'effet structurant de ces mobilités professionnelles, à l'échelle locale.

On observe que cette représentation des navettes de l'Isère en fonction d'un voisinage ordinal fait écho aux cartes présentées sur la Figure 7-1 (page 288), où le rôle de la distance parcourue par les flux est mis en évidence en utilisant un critère d'ordre sémiologique.

Plus généralement, la sélection des flux réalisée selon une métrique d'éloignement continue (D_{ij}) ou discrète (CK_{ij}) entraîne une réduction importante de la matrice.

Figure 7-16. Cartographie des flux de l'Isère (F_{ij}) \leftarrow [(CK $_{ij}$), k=4]

Nous allons maintenant envisager un troisième cas de prise en compte de l'espace dans la cartographie de flux : celui où la matrice de flux (F) peut être mise en relation avec une partition (P) de l'espace d'étude. Cela nous conduit à raisonner du point de vue de la proximité relative des (i) lieux les uns par rapport aux autres et de leurs influences respectives, devant aboutir à la formation de zones ou « régions » de nature économique, linguistique, culturelle, historique, diplomatique ...

7.3 La similarité des lieux : le critère d'appartenance territoriale [$F_{ij} \leftarrow (P_i = P_j)$]

Lorsque les unités spatiales concernées par l'analyse cartographique de flux sont de type surfacique – mais la solution vaut aussi pour des unités ponctuelles, si on peut les affecter à une maille et une seule – il est une mesure particulière d'éloignement (ou de proximité des lieux) non négligeable : la *relation d'appartenance* (politique ou culturelle, linguistique, économique, financière, ...) entre les (i) lieux, évoquée dans la section 1.3.3 (page 48). L'existence d'une telle relation traduit le fait que les zones d'origine et de destination sont proches non seulement parce qu'elles sont limitrophes (C_{ij}) ou voisines (CK_{ij}), mais parce qu'elles appartiennent à une même entité territoriale de niveau supérieur ($P_i = P_j$), c'est-à-dire elles partagent une caractéristique commune d'appartenance à une maille donnée.

Cette partition discrétisant l'espace des couples de lieux (i, j) en deux catégories binaires (appartient / n'appartient pas) à une même région donnée, « [...] implique que l'on peut définir des maillages soit à travers l'étude des attributs d'un ensemble d'éléments, soit à travers l'étude des relations entre les éléments d'un ensemble [...] le maillage est à la fois un réseau de relations et une partition » (Grasland C., 1997). La partition peut être endogène, c'est-à-dire déduite des flux ou exogène, liée à d'autres critères non liés aux flux.

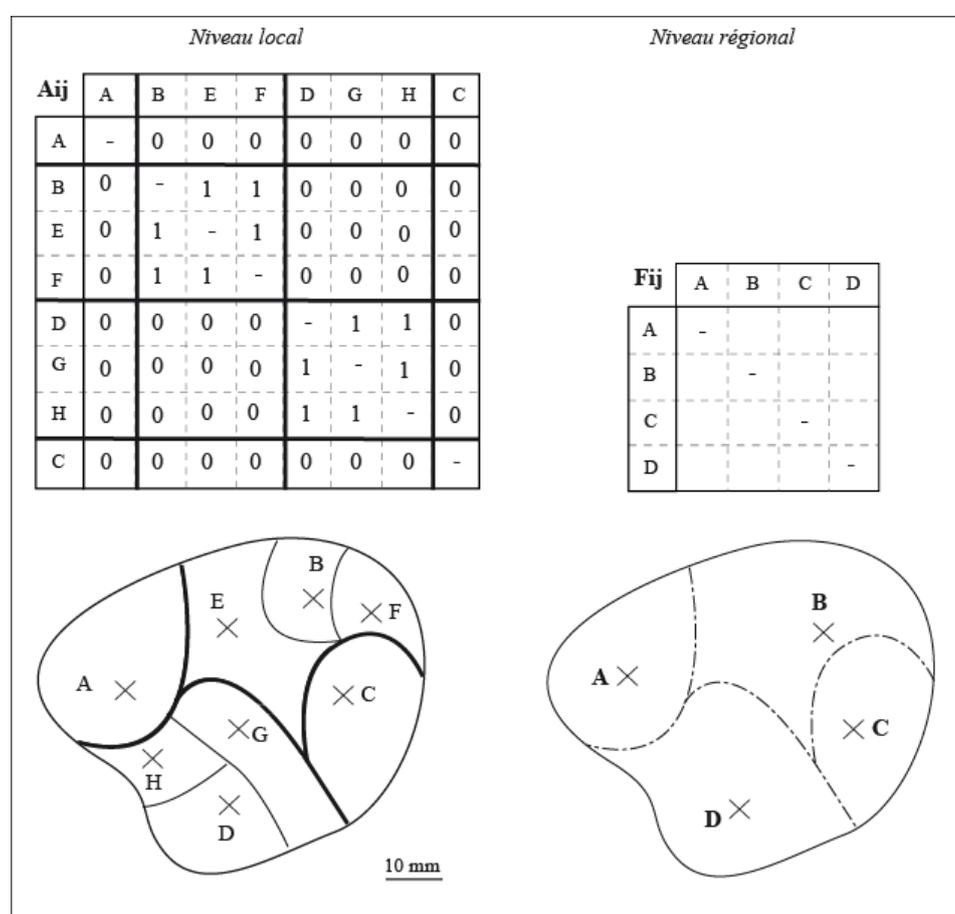
Etant donné que les partitions territoriales générées à partir des valeurs de flux posent un certain nombre de questions méthodologiques qu'il n'est pas envisageable de traiter ici, nous allons restreindre notre propos au cas des zonages structurels de l'espace des flux réalisés à partir de maillages hiérarchisés. Les conséquences d'une telle procédure conduisent à une régionalisation de l'espace des lieux. Son objectif n'est pas tant de transformer les données que de les présenter à une autre échelle spatiale (Tobler, 1981b : 9). Les méthodes correspondantes s'inscrivent par conséquent

dans la famille de solutions à la cartographie de flux liée à l'agrégation des données, certaines d'entre elles ayant d'ailleurs été proposées par W. Tobler (1981a, 1987, 1995).

L'importance de la question de la résolution du maillage de l'espace des flux nous amène à considérer que l'espace fait l'objet d'une partition territoriale structurale préexistante. Cette partition est déjà disponible pour l'analyse des flux, car elle a été élaborée indépendamment de la matrice de flux. Cela signifie que nous n'allons pas présenter de méthodes d'agrégation spatiale ou attributaire permettant de générer cette partition territoriale, ni discuter les objectifs de leur construction ; plusieurs partitions pouvant coexister sur un même territoire.

Considérons (A_{ij}), une matrice d'appartenance territoriale à des mailles de niveau supérieur à celui de l'échelle élémentaire d'observation des données de flux. Cette matrice est nécessairement bloc-diagonale. Sa construction ne pose pas de problème particulier (voir Figure 1-13, page 52). L'application de ce principe, sur la matrice d'exemple local en huit lieux, est présentée sur la Figure 7-17.

Figure 7-17. Illustration de la construction d'une matrice d'appartenance (A_{ij})

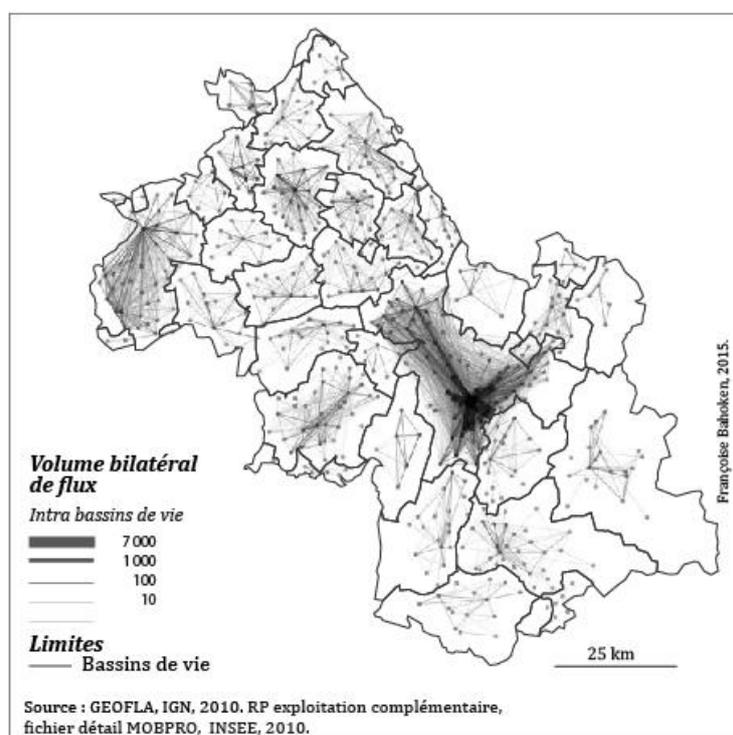


On voit bien ici le double intérêt de mobiliser (A_{ij}) : en agrégeant les unités spatiales élémentaires à un échelon supérieur, en fonction de leur appartenance commune, (A_{ij}) procure un découpage de (F_{ij}) en deux types de flux : des *flux intra régionaux* qui se produisent entre les unités élémentaires qui appartiennent à une même région – le terme région étant toujours pris ici au sens général d'agrégat d'unités élémentaires – et des *flux inter régionaux* qui se produisent entre des unités élémentaires qui appartiennent à des régions différentes.

Pour illustrer la procédure de construction de (Aij), nous allons considérer la matrice (Fij⁺) des flux de l'Isère en lien avec le découpage du département en bassins de vie défini par l'INSEE en 2012. Ce fichier « [...] constitue le plus petit territoire sur lequel les habitants ont accès aux équipements et services les plus courants. On délimite ses contours en plusieurs étapes. On définit tout d'abord un pôle de services comme une commune ou unité urbaine disposant d'au moins 16 des 31 équipements intermédiaires. Les zones d'influence de chaque pôle de services sont ensuite délimitées en regroupant les communes les plus proches, la proximité se mesurant en temps de trajet, par la route à heure creuse. Ainsi, pour chaque commune et pour chaque équipement non présent sur la commune, on détermine la commune la plus proche proposant cet équipement. Les équipements intermédiaires mais aussi les équipements de proximité sont pris en compte. » (INSEE, 2012)¹⁹⁴.

Le critère d'appartenance régionale caractérisant une forme particulière de similarité entre des zones (Aij=1), parce qu'elles appartiennent au même bassin de vie, la création de ce zonage passe ainsi par la fusion (ou l'agrégation attributaire) du fond de carte des communes de l'Isère, puis par la représentation des flux qui ne franchissent pas la limite régionale, comme illustré sur la Figure 7-18.

Figure 7-18. Cartographie des flux intercommunaux et intra bassins de vie de l'Isère



Sur la Figure 7-18, on observe la superposition des deux maillages (des communes et des bassins de vie) sur le plan géométrique. Les 533 communes iséroises sont regroupées en 38 bassins de vie de taille variables : certains d'entre eux sont en partie localisés sur des départements voisins et ne comportent qu'une commune située en Isère (Aoste, les Avenières et Romans-sur-Isère) et l'un d'entre eux est formé de 75 communes (Grenoble). Le nombre moyen de communes par bassin de vie est de 14 et la médiane de 11. Il est donc possible de représenter une sélection de flux qui concerne un bassin de vie en particulier ou l'ensemble des flux internes aux bassins de vie.

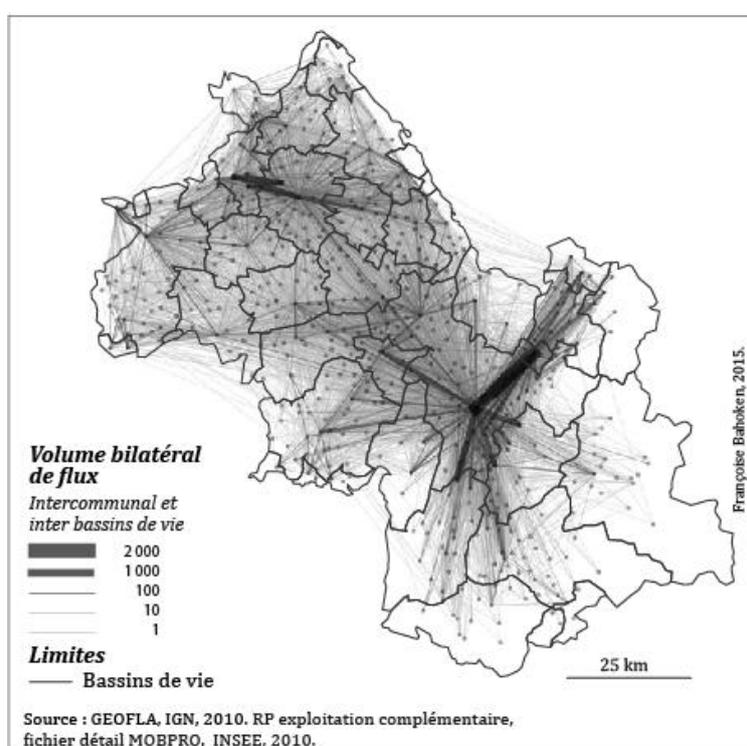
La Figure 7-18 présente les flux intra bassins, autrement dit les flux inter communaux qui n'ont pas franchi la limite des bassins de vie. Ces flux correspondent formellement au contenu de la diagonale

¹⁹⁴ Source : http://www.insee.fr/fr/themes/document.asp?ref_id=ip1425#encadre1

de la matrice. Cette figure met en évidence des sous-matrices qui correspondent à autant de sous-systèmes d'échanges ou migratoires fortement structurés. On observe à la fois une forte connectivité de ces graphes et une polarisation des flux vers la commune-centre du bassin. Elle traduit une relation de *continuité spatiale* (CSij) des flux entre les communes appartenant à chacune des régions, la *continuité spatiale* signifiant l'appartenance à un graphe connexe. Ces flux étant intra régionaux, il est possible de représenter des flux inter régionaux.

Tout en restant à l'échelle communale, on peut représenter des flux inter bassins soit à l'échelle communale, soit à celle des bassins. La carte de la Figure 7-19 représente les flux intercommunaux entre les bassins de vie, correspondant à l'inverse de la précédente. On observe que deux voire trois communes polarisent ces flux inter bassins : il s'agit en réalité d'une représentation de l'attractivité des pôles d'emplois départementaux (le découpage territorial en pôle d'emploi, également défini par l'INSEE, correspondant à une partition du département de l'Isère, en trois zones).

Figure 7-19. Cartographie des flux intercommunaux et inter bassins de vie de l'Isère

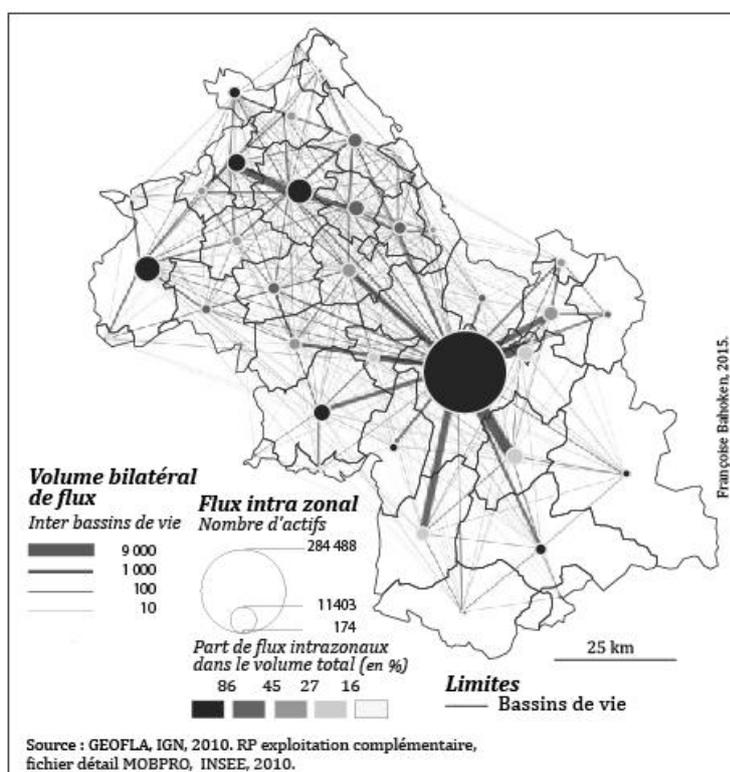
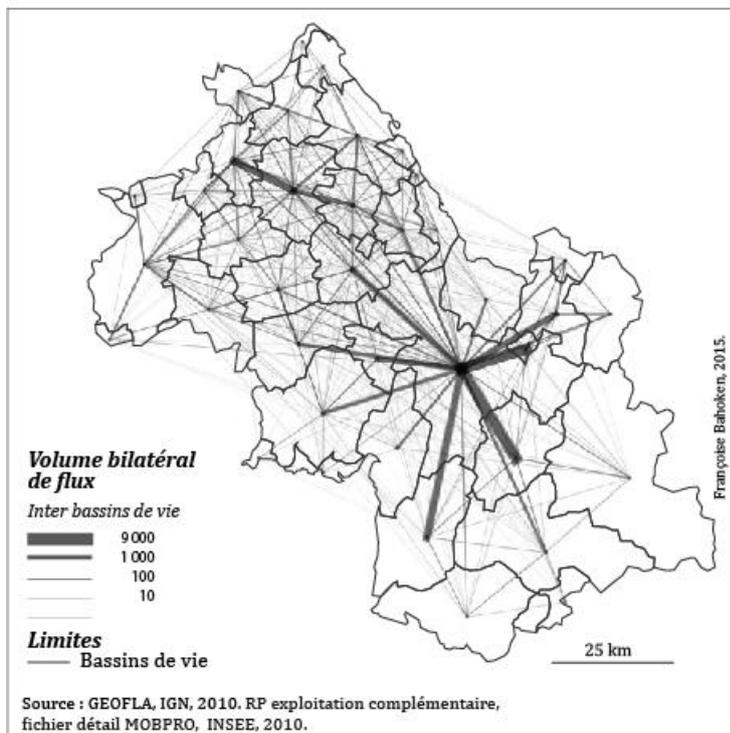


Une troisième possibilité consiste à simplifier les valeurs de flux inter bassins de vie. Pour cela, il est nécessaire de quitter l'échelle des communes pour observer la distribution des valeurs de ces flux à celle des bassins de vie, grâce à une procédure d'agrégation. Ce changement de résolution du maillage va conduire à un changement d'échelle d'observation : le passage de la commune au bassin de vie ; les valeurs de flux intercommunales seront, elles aussi, agrégées, ici par la fonction somme, à l'échelle des bassins.

La Figure 7-20 représente les volumes bilatéraux de flux inter bassins de vie. L'information locale n'étant plus disponible, elle ne sera pas représentée. Il est toutefois possible de la retrouver en cartographiant les volumes de flux intra zonaux, qui correspondent au volume total de flux échangé entre les communes qui composent chacun des bassins. Cette information est présentée sur la carte du bas de la Figure 7-20. Y a été ajoutée une coloration discrète des symboles proportionnels qui renseigne sur la part de ces flux intra zonaux dans le volume total de flux du bassin de vie.

On observe que le bassin de vie de Grenoble est le plus important en taille et en nombre de communes. Nombre des actifs résidents y travaillent : plus de 86% des mobilités professionnelles grenobloises ont lieu au sein même du bassin. A l'inverse, les bassins de vie montagnards voient leurs actifs en grande majorité descendre dans le bassin de vie grenoblois pour aller y travailler.

Figure 7-20. Volumes bilatéraux de flux inter bassins de vie de l'Isère



Réduire (Fij) en fonction de (Aij) conduit donc à une forme de sélection des flux en fonction d'un échelon territorial particulier : l'échelle locale ou régionale. Disposer ainsi de cartes de flux à différents niveaux d'agrégation ouvre la voie vers les analyses multi-niveaux, dans la mesure où les flux peuvent être simultanément cartographiés selon qu'ils sont interrégionaux (et ne concernent que l'échelle régionale) et/ou intra-régionaux.

Cette possibilité de cartographie de flux qui considère différents niveaux géographiques fait écho aux travaux réalisés par L. Chapelon (1996), A. l'Hostis et P. Mathis (1996) sur les réseaux de transports. Les méthodes mobilisées par ces auteurs sont fondées sur les notions de « *zooms nodaux de niveau n-1 ou multi nodaux* », elles conduisent à des représentations des réseaux qu'ils qualifient de quasi-fractales. En effet, les cartes obtenues correspondent à « [...] *la représentation simultanée de réseaux localisés à des échelles géographiques différentes. Chacune de ces échelles renvoyant à un niveau d'organisation spatiale spécifique* » (Chapelon, 1996 : 29-30). Appliquées aux réseaux de transport (formellement à des graphes planaires valués ou matrices de type (Lij)), nous faisons l'hypothèse – que nous n'investiguerons pas dans la thèse – que ces méthodes pourraient être étendues aux matrices de flux, de manière à enrichir leur cartographie par la prise en compte du (ou des) contexte(s) d'expression des flux.

Si elle répond bien à l'objectif de simplification des matrices de flux denses, la procédure générale d'agrégation des valeurs de flux a pu être controversée par certains auteurs, car elle conduit *in fine* à dégrader l'information, voire à changer de concept par le changement d'échelle opéré. D'après C. Cauvin et H. Reymond (1991), une manière « *simple mais discutable* » qui vient d'emblée à l'esprit vise à réduire la quantité de flux à représenter en effectuant une « [...] *réduction par combinaison, regroupement des lieux adjacents, c'est-à-dire contigus. On diminue ainsi le nombre des lieux entre lesquels se font les échanges, mais dans le même temps, il est bien évident que l'on 'gomme' des détails de la représentation des mouvements* » (Cauvin C., Reymond H, 1991 : 480).

7.4 Validation empirique : l'exemple de flux aériens internationaux de passagers

Cette section porte sur la validation empirique des propositions de cartographie de flux prenant en compte une variable spatiale liée à l'éloignement des lieux : la distance géographique (Dij), la proximité (Cij) ou (CKij) ou l'appartenance territoriale commune (Pi=Pj). Nous mobilisons pour ce faire, une matrice de flux mondiaux s'exprimant entre des unités spatiales ponctuelles. Il s'agit de flux de passagers internationaux transportés par voie aérienne, pour l'année 2000 ; ces informations sont issues de la base de données de l'Organisation de l'Aviation Civile Internationale (OACI). Elles ont été notamment utilisées par N. Cattani (2004) dans un article publié dans la revue *Flux* (mobilisant les valeurs des années 1981, 1990 et 2000) auquel nous ferons référence dans un premier temps.

Nous présentons dans une première sous-section les données disponibles dans cette matrice. Nous la segmentons ensuite en fonction de critères de distances continues (Dij) et discrète (Cij=1) dans la seconde section, puis en fonction d'une acception de la similarité des pays.

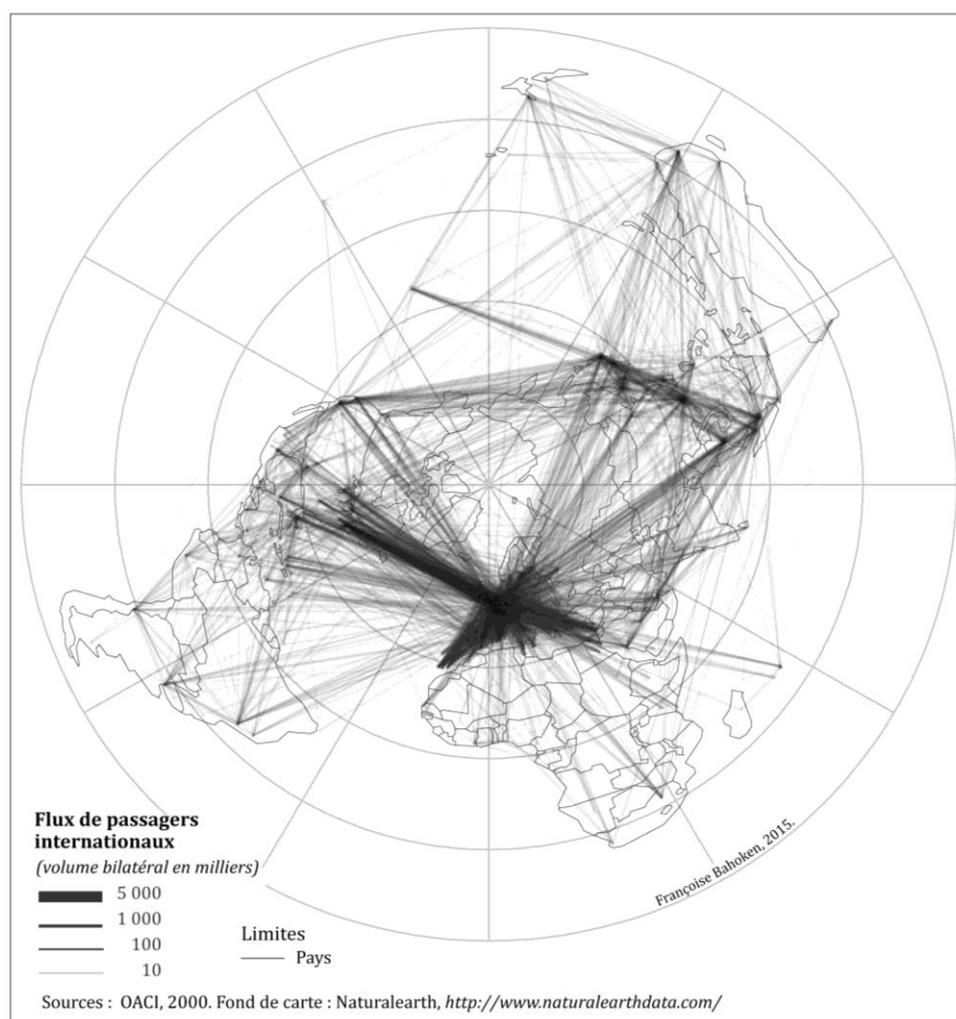
7.4.1 Présentation des données

La matrice de l'OACI décrit les flux de passagers internationaux entre des aéroports mondiaux. L'intérêt du recours à cette donnée réside dans le fait qu'elle diffère à la fois du cas de l'Isère et de celui de flux mondiaux de marchandises. En effet, ces flux aériens concernent des populations humaines dont on connaît le mode de transport. Ils sont observés au niveau de l'aéroport, ce qui correspond à un mode d'implantation ponctuelle et non surfacique comme dans les exemples précédents. Cette particularité ouvre de fait des possibilités intéressantes de transformation des

données (d'une implantation ponctuelle à une implantation surfacique) et d'agrégation d'un niveau à un autre, d'une échelle géographique à une autre. De plus, la disponibilité de ces flux à l'échelle mondiale n'interdit pas leur description à l'échelle nationale : de l'ensemble des flux entre des aéroports appartenant à un même pays ou bien à l'échelle internationale des aéroports (d'un voire de plusieurs pays) qui émettent ou reçoivent des vols internationaux, si l'on prend en compte le critère du le franchissement d'une frontière nationale.

Cette matrice compte 10 628 liaisons disponibles de plus de 10 000 passagers, entre 436*436 villes aéroportuaires qui sont décrites par leurs coordonnées (x, y) et par leur appartenance nationale. La matrice étant symétrique, elle correspond formellement à : $(F_{ij}^+) > 10\,000$ passagers. La cartographie de l'information disponible dans cette matrice est présentée sur la Figure 7-21, dans le formalisme du trait droit.

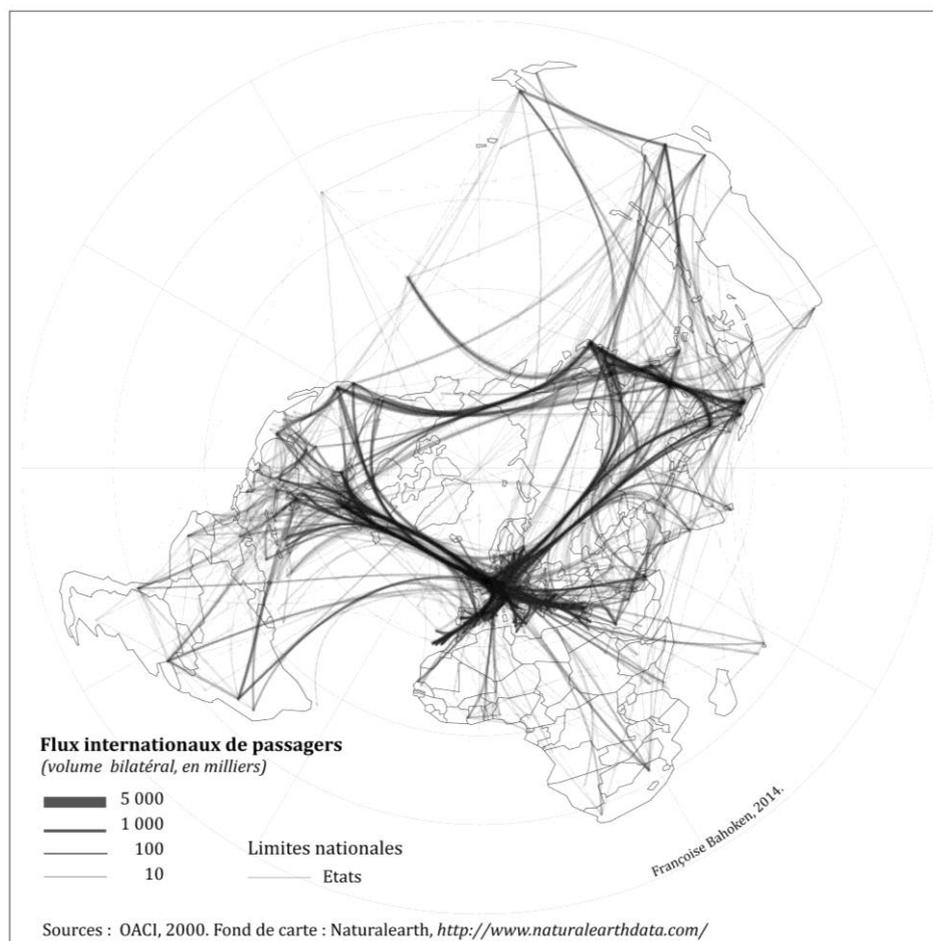
Figure 7-21. La vision polaire des flux aériens internationaux de passagers en 2000 (version 1)



Cette Figure 7-21 est peu lisible, en raison de la manifestation d'un *effet spaghetti*. Ce constat justifie les procédures de restriction appliquées par N. Cattani à ces données (suppression des flux intra européens, représentation des seuls 100 premiers flux, ...) lors du traitement des données numériques. Il est désormais possible de réduire la complexité graphique de cette figure en raisonnant au niveau graphique. Cependant, en l'absence d'informations sur le tracé exact de ces routes aériennes, nous proposons de représenter la masse d'information de ces flux, en les fusionnant graphiquement, selon la

méthode de l'*edge bundling* de Holten (2005). Cela nous permet de disposer d'une vision globale et lisible de l'ensemble de ces flux aériens internationaux mondiaux de passagers. La carte correspondante est présentée sur la Figure 7-22.

Figure 7-22. La vision polaire des flux aériens internationaux de passagers en 2000 (version 2)



Il convient de noter que le passage entre les deux figures précédentes est purement formel et dépend de nos choix, tous permettant de décrire plus ou moins finement la structure des principaux flux aériens entre les pôles. Comme l'indique N. Cattan, « [...] *les plus forts courants de trafic aérien international [qui] arrivent certaines villes des pays « du sud » au système mondial des échanges. Il s'agit par exemple des liens entre Dubaï et Londres et entre Buenos-Aires, Miami et Madrid* » (Cattan, 2004 : 34) sont bien conservés. La figure permet de rendre compte des rapports de force internationaux que l'attractivité touristique de certaine villes traduit. Elle répond aussi bien aux impératifs de perception des grands centres économiques mondiaux dont parle N. Cattan, puisque les principaux centres ne sont plus alignés dans l'hémisphère nord, mais bien répartis au cœur de la planche. Cette représentation sied, d'ailleurs, également mieux à la perception de la position des villes du sud à l'échelle mondiale.

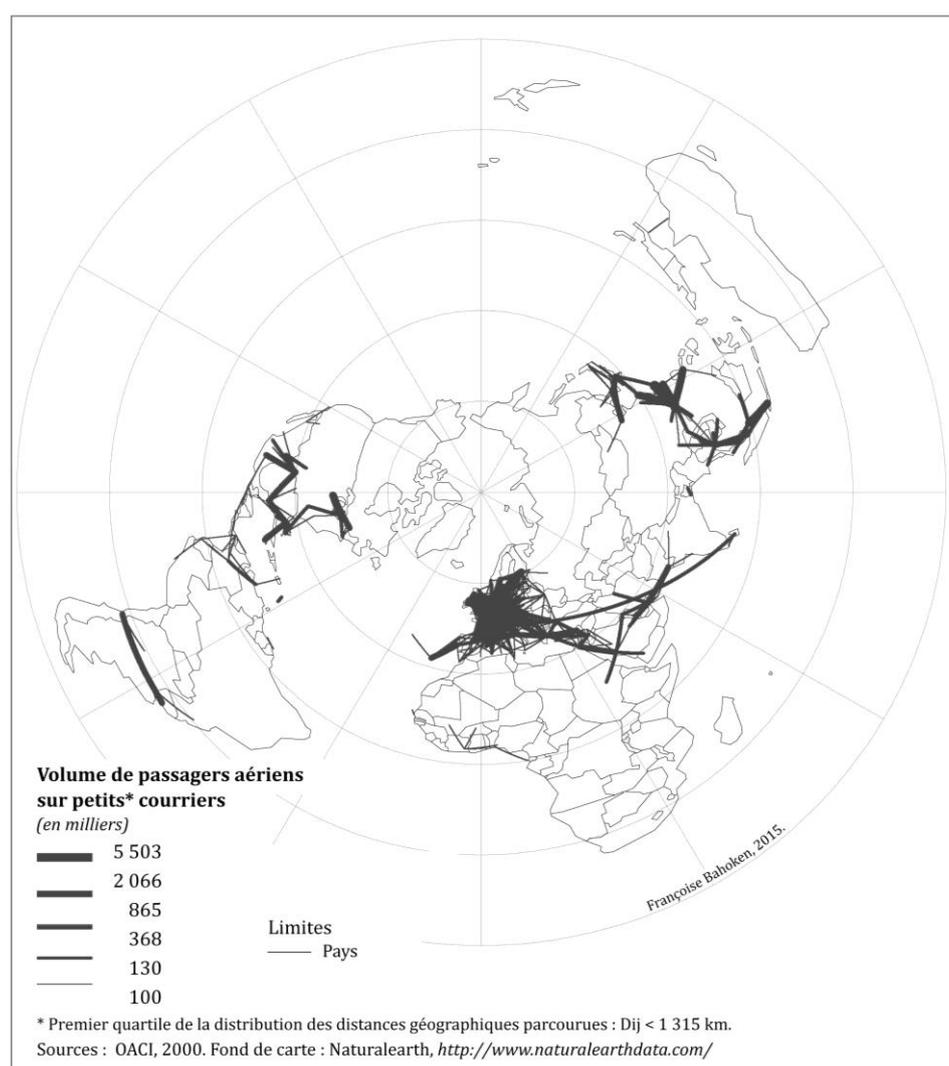
Dans la section suivante, nous allons intégrer l'espace de trois manières, ce qui nous conduira à nuancer les motifs de la carte présentée sur la Figure 7-21. Nous mobiliserons tour à tour les matrices de type (Dij), (Cij) et (Aij).

Afin de pouvoir disposer de cartes comparables, nous faisons le choix de cartographier les flux internationaux de passagers à courte distance ou locaux, ces derniers correspondant aux avions de ligne de type petit courrier¹⁹⁵.

7.4.2 Cartographie des vols aériens locaux : le « petit courrier » ($F_{ij} \leftarrow (D_{ij} < Q1)$)

La cartographie des flux internationaux de passagers à courte distance selon une métrique continue, la distance géographique, par exemple, passe par une réduction de la matrice (F_{ij}) en fonction de (D_{ij}) : ($F_{ij} \leftarrow (D_{ij} < Q1)$). La carte de la Figure 7-23 décrit ainsi les flux aériens dont la distance parcourue est inférieure à 1 315 km.

Figure 7-23. Cartographie du volume de flux de passagers internationaux en petits courriers



Ces flux peuvent être analysés soit en termes de fréquence de relation (on observe de nombreux flux sur de courtes distances au sein des pays de l'Union européenne des quinze (UE-15) correspondant à la date des données, où le maillage territorial est fin), soit en termes d'effectifs de passagers. C'est cette seconde possibilité qui est envisagée ici, la première sera examinée dans le chapitre 8.

¹⁹⁵ Dans l'aviation civile, on distingue généralement trois types d'avions de ligne que sont les petits courriers, les moyens et les longs courriers, ces derniers étant par ailleurs ds gros porteurs (transportant les effectifs de passages les plus importants sur de longues distances).

L'objectif étant de visualiser la courte distance, nous avons supposé qu'elle correspond au premier quartile (Q1) de (Dij), c'est-à-dire à des flux parcourant de moins de 1 315 km. Le Tableau 7-2 résume la matrice (Dij) des distances kilométriques orthodromiques parcourues par les flux internationaux de passagers pour l'année 2000.

Tableau 7-2. Résumé statistique de la matrice (Dij) des liaisons aériennes (Lij)

(Dij)	Min	Q1	Q2	Q3	Max	Moyenne
Distance kilométrique	47	1 315	2 188	3 755	18 377	3 152,1

Source : OACI, 2000

La carte de la Figure 7-23 décrit ainsi les vols dits de « petits courriers » qui concernent les flux aériens effectués par de petits porteurs, des avions de ligne présentant une capacité de charge peu importante, caractéristiques de trajets effectués sur de courtes distances. Ces flux concernent aussi des vols de fréquences élevées qui concernent, dans l'ensemble, une population importante que les vols long courrier. Autrement dit, ces lignes concernent des flux présentant de faibles quantités de passagers sur de courtes distances, lorsqu'ils sont pris individuellement, mais dans des fréquences élevées, d'où cet *effet de surcharge graphique* qui peut apparaître important au sein de l'UE-15, alors que la *couverture spatiale* est globalement faible.

La seconde possibilité de représentation de la courte distance consiste à raisonner selon une métrique discrète, en représentant les flux qui se produisent entre des pays contigus ou limitrophes ($C_{ij}=1$), cette possibilité est présentée dans la sous-section suivante.

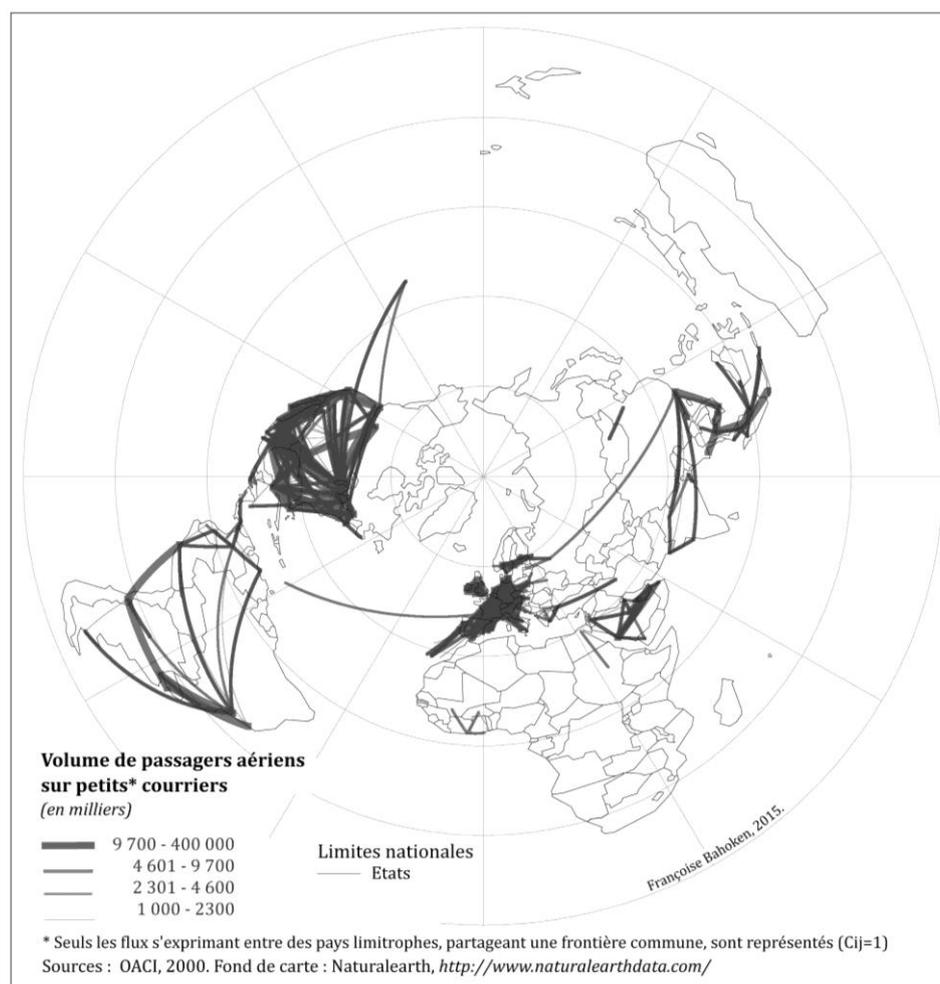
7.4.3 Cartographie des flux aériens entre pays limitrophes : le petit courrier entre voisins (Fij) ← ($C_{ij}=1$)

La cartographie des flux à courte distance varie nettement en fonction de l'échelle de représentation. A l'échelle mondiale, la cartographie des flux réalisés entre des pays voisins renvoie à deux notions distinctes : celle de voisinage ($C_{ij}=1$), comme à l'échelle locale, et celle d'appartenance régionale. Dans le premier cas, le critère mobilisé est la contiguïté ordinaire qui entraîne la formation d'un graphe connexe décrivant les relations entre des pays qui appartiennent à un même continent, comme nous avons pu le voir dans la section 7.3 (page 306). La Figure 7-24 décrit ainsi les petits courriers desservant des pays limitrophes.

Le voisinage mondial concerne les flux qui sont d'une part, réalisés entre des états qui partagent une frontière nationale commune, donc les pays limitrophes¹⁹⁶; d'autre part, les flux émis entre des pays qui appartiennent au même continent, mais qui ne sont pas forcément voisins.

Dans le premier cas (voir Figure 7-24), la carte met en évidence des flux locaux, souvent qualifiés de régionaux, qui correspondent aux petits courriers tandis que, dans le second cas, elle décrira plutôt des sous-régions formées des relations connexes. Ces dernières correspondent aussi à des flux d'ampleur régionale, à des moyens courriers car elles sont réalisées à un niveau subcontinental (il s'agit de flux intracontinentaux). La différence entre les deux cartes réside dans l'échelle d'observation de la relation de voisinage : au niveau national (« Je cartographie les flux entre des pays voisins ou limitrophes »), correspondant aux petits courriers et dans le second, la carte est celle l'échelle continentale, par exemple : (« Je cartographie les flux qui se produisent entre des pays qui appartiennent au même continent, mais qui ne sont pas nécessairement voisins »).

¹⁹⁶ Il convient de noter qu'en raison de pays multisitués, tels que la France ou encore l'Espagne, la continuité spatiale d'un pays n'est pas toujours assurée avec ce type de flux.

Figure 7-24. Cartographie du volume de flux de passagers internationaux entre pays voisins

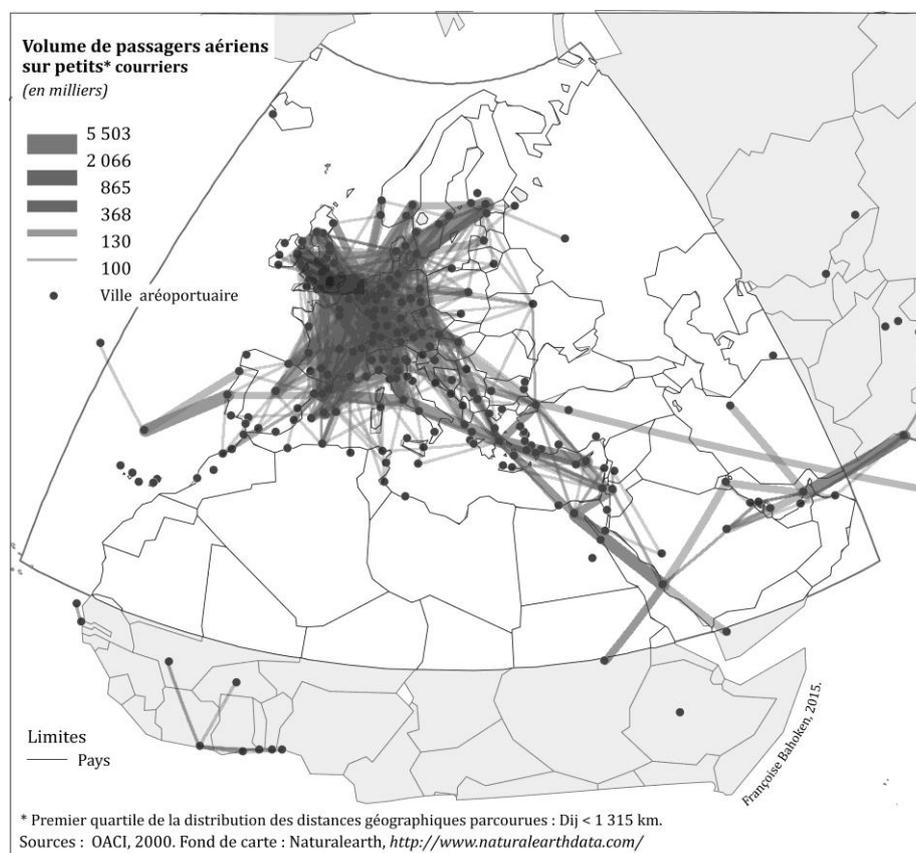
La comparaison de la carte de la Figure 7-24 avec celle de la Figure 7-23 met en évidence l'importance du choix du critère de sélection des flux selon l'espace. Elles décrivent toutes deux des mobilités locales, sous-régionales, aux motifs sensiblement équivalents. Sur le plan graphique, *l'effet de couverture spatiale* est absent, tandis que *l'effet de surcharge graphique* est important, mais localement, lorsque le maillage territorial est dense. En effet, la fréquence des flux apparaît forte au sein des pays de l'UE-15 et de la côte est des Etats-Unis dont les populations, plus aisées, sont plus à même de se déplacer par voie aérienne sur des petits et moyens courriers. Nous pouvons illustrer cela avec le cas des flux nord-américains, réalisés entre des pays dont la superficie est importante.

Lorsque la sélection est effectuée sur un critère de distance kilométrique (voir Figure 7-23), ces flux nord-américains apparaissent très localisés tandis que, à l'inverse, le critère de voisinage (voir Figure 7-24) montre l'ampleur des échanges nord-américains correspondant aux moyens courriers. Ce résultat souligne l'importance de la définition du périmètre et du niveau d'observation concerné, comme nous avons pu le montrer dans la section 7.3. Une discussion sur le cas spécifique de l'échelle mondiale est d'ailleurs présentée par G. Van Hamme et C. Grasland (2011), à propos du découpage du monde en dix zones fonctionnelles.

Compte tenu de la spécificité de ces moyens courriers, il est ainsi possible de raisonner au niveau mondial en agrégeant localement les flux qui sont très fréquents. Nous allons focaliser l'analyse sur l'UE-15 qui correspondra au périmètre extérieur de la zone d'étude. Nous examinerons en particulier l'espace euro-méditerranéen (EUROMED), tel qu'il est défini dans le projet ESPON 1. Le recours à

l'EUROMED conduit ainsi à changer d'échelle et à focaliser l'analyse sur une région du monde (l'Europe) et non sur l'ensemble, comme illustré sur la Figure 7-25.

Figure 7-25. Cartographie du volume de flux de passagers internationaux, sur l'espace euro méditerranéen, en petits courriers



Cette figure permet de décrire la fréquence (importante) des flux de passagers aériens entre les villes européennes qui dépend du nombre (important) de villes concernées par l'analyse, sur un territoire plutôt restreint, par rapport à d'autres parties du monde (il s'agit des villes répondant au critère de sélection de l'OACI et accueillant des liaisons aériennes de plus de 10 000 passagers). La focalisation sur l'EUROMED ne permet pas de visualiser plus clairement les liaisons intra européennes pour des raisons liées à la résolution du maillage. En effet, « *C'est en partie un effet d'optique qui vient du fait que l'Europe est un continent politiquement très morcelé, ce qui multiplie d'autant les liens bilatéraux. Beaucoup de pays de cette région ont des relations commerciales privilégiées avec des pays d'autres régions du monde* » (Richard, Zanin, 2009 : 143).

Cette focalisation sur l'EUROMED étant, en outre, difficile à généraliser – elle ne correspond pas, à notre connaissance, à l'échelon d'un découpage harmonisé à l'échelle du globe –, nous allons l'abandonner momentanément. Dans la prochaine section, nous allons revenir à une échelle d'analyse mondiale et pour cela, nous utiliserons une partition territoriale couvrant l'ensemble mondial.

7.4.4 Cartographie des flux aériens (inter)régionaux : les moyen et long courriers (F_{ij}) $\leftarrow (P_i=P_j)$

Les flux internationaux de passagers aériens peuvent être agrégés selon différents découpages du monde, sous réserve qu'ils couvrent l'ensemble des données et de préférence, qu'ils soient harmonisés. Les données de l'OACI étant disponibles à l'échelle des villes, il est possible d'utiliser un premier niveau d'agrégation qui serait le pays. Cependant, celui-ci n'est pas suffisamment important

pour notre démonstration. Plusieurs possibilités de découpage du monde en grands ensembles supranationaux sont en outre envisageables et le choix est vaste quant aux critères de découpage à retenir (des régions naturelles ou non, des limites maritimes, continentales ou mixtes) et à l'approche mobilisée (politique, économique, symbolique, culturelle, environnementale, ...).

Sans rentrer dans les discussions portant sur le choix du découpage qui serait le plus adapté aux données sur les flux aériens que nous mobilisons, nous proposons de raisonner selon une approche macro économique dont l'objectif serait, par exemple, de comparer le niveau de développement et l'importance des flux à l'échelle mondiale. Parmi les partitions économiques du monde existantes et compte tenu de la date des données (2000), nous faisons le choix du découpage du monde réalisé en 2006, dans le cadre du projet *Europe in the World* (EIW)¹⁹⁷.

Le *World Unified Territorial System* (WUTS), inspiré de la nomenclature en NUTS¹⁹⁸ (*Nomenclature of Territorial Units for Statistics*) d'EUROSTATS, est un zonage territorial harmonisé à l'échelle mondiale. Il est composé de cinq échelons hiérarchiques emboîtés, qui sont définis de la manière suivante : le niveau WUTS-5 correspond aux limites nationales en 168 pays, le WUTS-4 à un découpage du monde en 12 régions, le WUTS-3 en 17 régions, le WUTS-2 en 7 macros régions, le WUTS-1 en 3 régions globales et le WUTS-1 en une région mondiale unique. Nous faisons le choix du niveau WUTS-2.

La Figure 7-26 présente le niveau WUTS-2, agrégeant le monde en 7 macros régions.

D'après le rapport du projet EIW, l'un des critères qui préside à la construction du niveau WUTS2, c'est-à-dire de ce regroupement d'Etats ou de groupes d'Etats, est le degré de développement économique. Cela explique, par exemple, la division de l'Asie et du Pacifique en deux régions. En fonction des parties du monde, le degré de proximité culturelle ou historique est également pris en compte, comme par exemple dans le cas de l'Amérique latine. Il convient enfin de garder en mémoire le fait que le découpage en WUTS a été élaboré dans le cadre des projets ESPON « [...] conduit à identifier des régions du monde d'intérêt européen qui pourraient être considérées comme non pertinentes pour d'autres parties du monde ». Le zonage du WUTS-2 présente ainsi l'intérêt de présenter des surfaces globalement homogènes. « En termes de taille à l'échelle mondiale, les sept régions macro sont relativement égales en termes de leur surface (plus ou moins 10 à 20% du monde pour chaque) ce qui est intéressant pour les seules raisons cartographiques ». (ESPON 3-4-1, Report WUTS : 83).

¹⁹⁷ La base de données EIW, ses métadonnées et la partie du rapport de la phase I du projet ESPON sont disponibles librement, à l'adresse suivante : <http://database.espon.eu/db2/resource?idCat=44> [vérifiée le 03/04/2015].

¹⁹⁸ La base de données du découpage de l'Europe en NUTS est disponible à l'adresse suivante : <http://ec.europa.eu/eurostat/web/nuts/overview> [lien vérifié le 03/04/2015].

Figure 7-26. Le découpage du monde en sept régions du niveau WUTS-2



Nous interprétons ainsi ce niveau WUTS-2 comme matrice d'appartenance à une macro région économique (A_{ij}) des différents pays d'origine et/ou de destination des flux aériens. Il est dès lors possible de cartographier ces flux de passagers à l'échelle intra ou inter régionale, en conservant ou non le niveau local de l'aéroport. Pour cela, il est nécessaire de construire au préalable la matrice des valeurs de flux interrégionaux équivalente au fond de carte des WUTS-2 : elle résulte de l'agrégation par la fonction somme, des valeurs de flux internationaux.

La matrice présentée sur le Tableau 7-3 a été réalisée à partir des données brutes de l'OACI (2000), hormis la condition de franchissement d'une limite nationale qui a été conservée.

Tableau 7-3. La matrice de flux inter macro régionaux de passagers aériens

		W11	W12	W13	W21	W22	W31	W32
Europe et Asie du nord	W11	521 747	30 517	8 636	61 704	10 104	12 096	14 830
Méditerranée du sud et Asie de l'ouest	W12	30 517	11 628	631	1 672	0	5 565	2 116
Afrique subsaharienne	W13	8 636	631	1 172	354	0	257	524
Amerique du nord	W21	61 704	1 672	354	48 084	22 618	3 466	23 300
Amérique latine	W22	10 104	0	0	22 618	13 368	0	292
Asie du sud et de l'est	W31	12 096	5 565	257	3 466	0	26 807	41 319
Pacifique occidental	W32	14 830	2 116	524	23 300	292	41 319	84 479

Valeurs arrondies, en milliers de passagers.

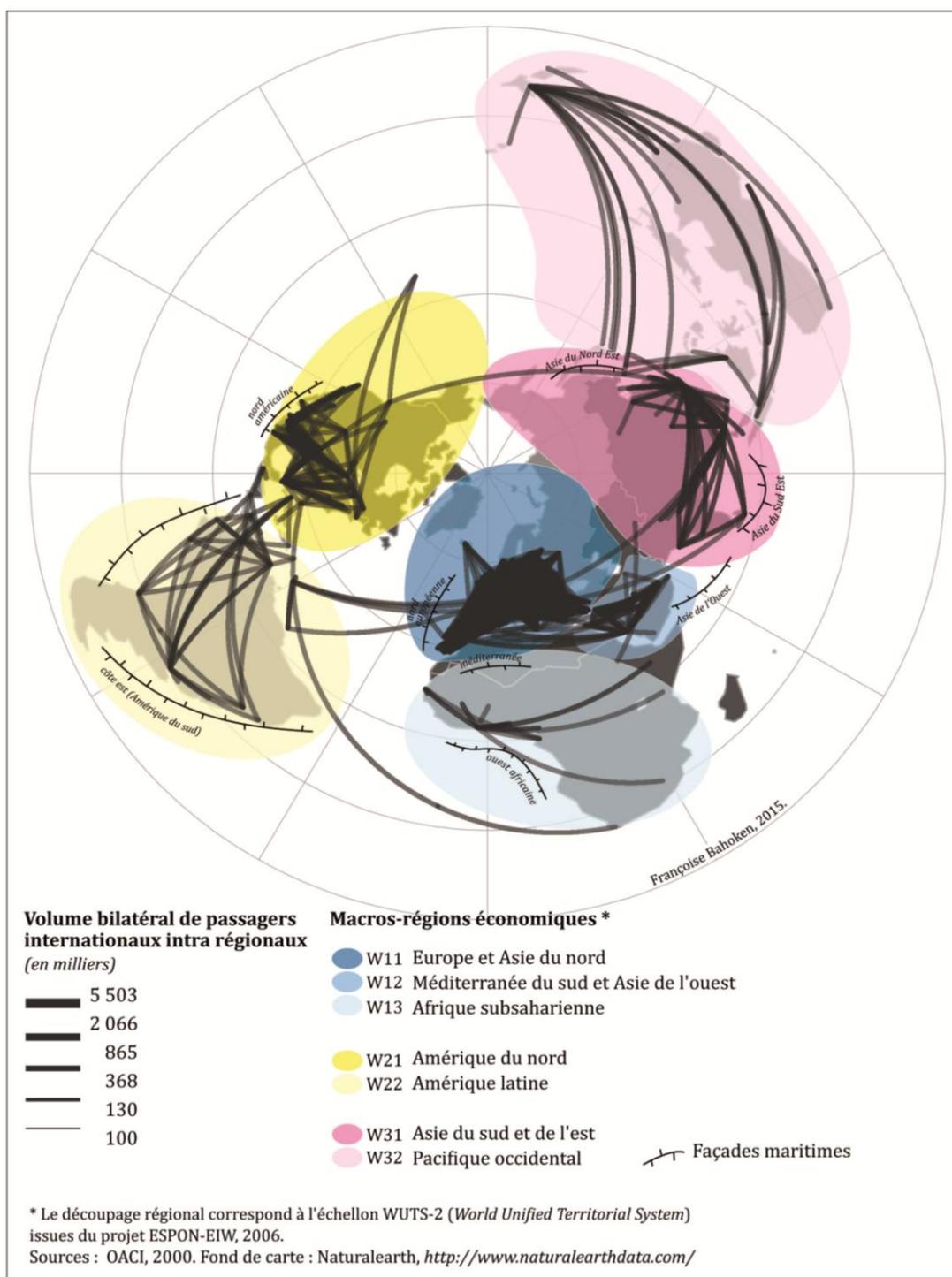
Source : OACI, 2000.

Cette matrice est *symétrique* (F_{ij}^+), tout comme celle des flux internationaux inter aéroports. Sa diagonale n'est cependant pas vide : elle est renseignée par le total des flux intra régionaux, de ceux qui n'ont pas franchi la limite d'une macro-région (des flux inter aéroports appartenant à une même macro-région). Etant donné que nous disposons de la matrice de ces flux élémentaires, il est possible de cartographier ces flux intrarégionaux à différentes échelles.

La Figure 7-27 décrit les flux cantonnés au périmètre des macro-régions. Sont également représentées les différentes macro-régions de niveau WUTS-2 par des enveloppes dont la teinte fait référence à celle de la Figure 7-26, ainsi que les principales façades maritimes.

On observe que certains flux entre Etats, réalisés à l'intérieur des macros régions, traversent des frontières maritimes. Ce phénomène n'est pas anormal : il concerne les régions dont les Etats possèdent des territoires outre-mer ou qui sont multi situés. C'est le cas, par exemple, des flux de passagers aériens au départ de Pointe-à-Pitre, en Guadeloupe, à destination de Cayenne, en Guyane, respectivement à un DOM et un TOM français, tous situés dans le bassin des Caraïbes et qui appartiennent à la région W11-Europe et Asie du nord et non Amérique Latine. Ces flux ne s'exprimant pas sur un territoire continu, ils franchissent deux frontières : une frontière continentale (l'Europe métropolitaine) et une frontière maritime (l'océan Atlantique). De la même façon, les flux qui desservent l'île d'Hawaï depuis les villes nord-américaines de Houston ou encore de Vancouver franchissent trois limites : régionale, continentale et maritime. Cela s'explique aussi par la méthode sous-jacente à l'élaboration de ces régions, qui est attributaire et non spatiale. Elle n'intègre de ce fait pas de critère de contiguïté ou de continuité spatiale.

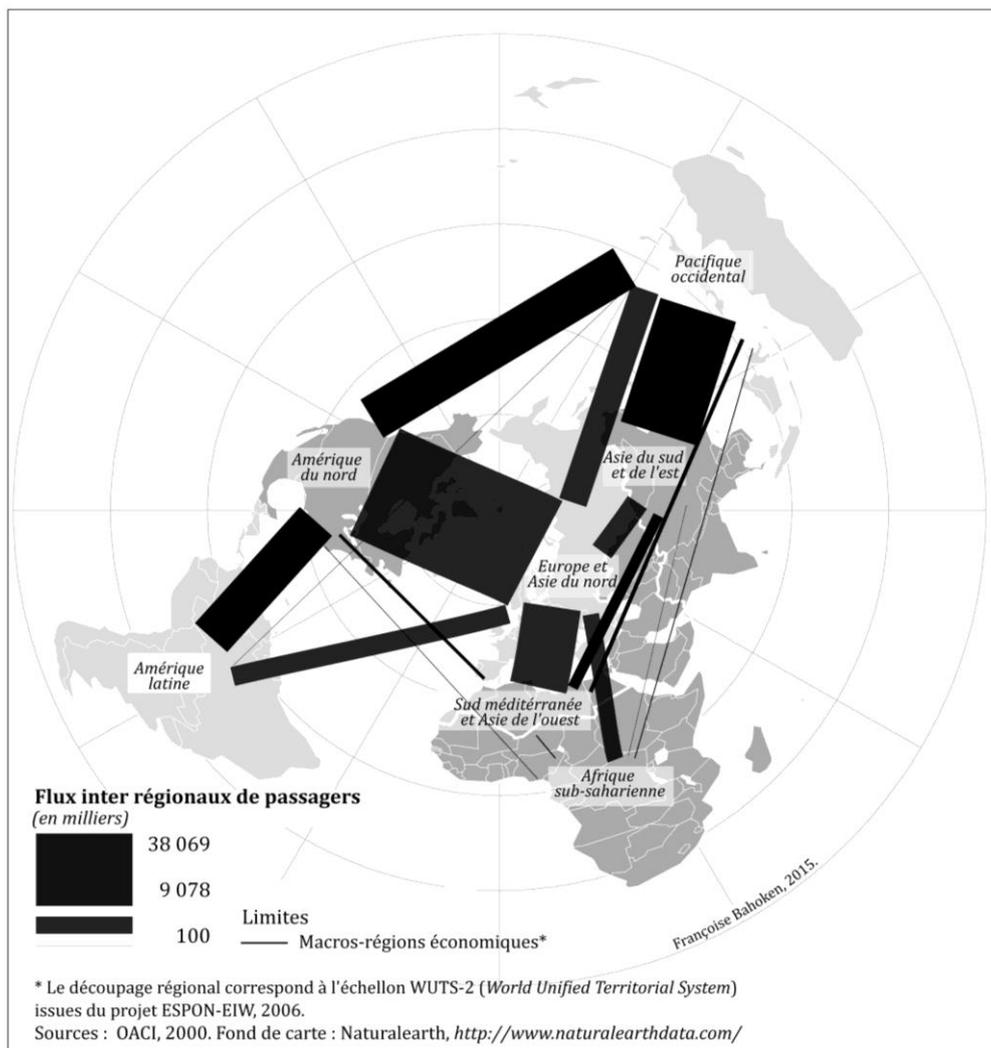
Figure 7-27. Cartographie du volume de flux inter macro régionaux mondiaux de passagers aériens



Il est également possible de ne cartographier que les flux inter-Etats qui franchissent la limite régionale, en tenant compte ou non de la continuité spatiale. Dans ce cas, la carte mettra alors en évidence les villes qui permettent de passer d'une région à une autre, alors que celle de la Figure 7-27 valorise les villes qui polarisent les flux au sein d'une même région. Sur la côté ouest-africaine par exemple, il s'agit, par exemple, de la ville d'Abidjan, en Côte-d'Ivoire et en Amérique Latine, de Santiago du Chili, de Buenos-Aires, de São Paulo ou encore de Caracas, au Venezuela.

Enfin, une dernière possibilité consiste à changer d'échelle, en représentant les flux sous une forme agrégée à l'échelle de ces macro-régions. La figure ci-dessous décrit les flux internationaux hors diagonale, entre les sept macros-régions économiques.

Figure 7-28. Cartographie du volume de flux internationaux inter régionaux de passagers aériens



Conclusion du chapitre

L'introduction de l'espace dans la procédure de cartographie de flux consiste à réduire la matrice de flux (F_{ij}) en fonction d'une matrice externe, caractérisant la distance géographique entre les lieux, (D_{ij}) par exemple. Si la construction de cette matrice est toujours intrinsèque aux données – sa mesure est directement déduite des coordonnées géographiques des lieux d'origine et de destination – elle n'en demeure pas moins exogène aux valeurs de flux. En effet, la cartographie des flux réalisée en fonction de classes de distance nécessite de mobiliser deux tableaux, le premier (D_{ij}) servant à réduire le second (F_{ij}), qui fait l'objet de la représentation.

En réponse à *l'effet de profondeur*, la matrice d'éloignement conduit à isoler les flux selon qu'ils sont locaux (et s'expriment entre des zones qui sont voisines, limitrophes ou proches), c'est-à-dire à courte distance, de ceux qui sont lointains et cela, sans qu'il faille se soucier outre mesure de la valeur du seuil qui sépare ces deux grandes catégories de flux. Ce résultat n'est pas anodin car, sur le plan conceptuel, les flux à courte distance diffèrent des flux à longue distance. Leur représentation cartographique n'est d'ailleurs pas similaire : les premiers peuvent être représentés sans ambiguïté à l'aide d'un trait droit, la distance cartographique correspondant globalement à la distance géographique, ce qui n'est pas le cas des flux à longue distance, comme nous l'avons vu dans le Chapitre 6. En réalité, c'est à mesure que l'on s'éloigne de l'échelle locale, que le flux prend de la distance, que le mouvement sous-jacent est moins lisible.

Le cas spécifique de cartographie des flux selon un critère de distance géographique continue ($D_{ij} < Q1$) ou de distance de proximité de voisinage ($C_{ij} = 1$) revêt toutefois un caractère particulier. En effet, il entraîne une réduction significative de la matrice de flux qui autorise la cartographie de l'ensemble de l'information répondant au critère d'éloignement (la matrice est suffisamment creuse pour être cartographiée directement). L'intérêt de ces solutions est qu'elles conduisent à réduire la partie de *l'effet spaghetti* qui est liée à la *couverture spatiale*. Les figurés représentés correspondent aux flux qui parcourent une certaine classe de distance, qu'elle soit réelle ($D_{ij} < Q1$), topologique ($C_{ij} = 1$) ou d'appartenance territoriale ($A_{ij} = 1$) à la même région, elles sont toutes caractéristiques d'une mobilité locale.

Le critère appliqué pour cartographier ces flux locaux, en fonction d'une acception de l'éloignement des lieux, est externe aux données numériques. Il mobilise une matrice (D_{ij}) – certes parfois élaborée à partir des coordonnées des lieux d'origine et de destination – pour réduire (F_{ij}) dont les valeurs sont représentées sur une carte. Mais il est possible de réduire la matrice (F_{ij}) sans recourir à un tableau externe, en appliquant un critère (α) en dessous duquel les flux ne seront pas représentés : $(F_{ij}) > \alpha$, ce critère pouvant agir à un niveau global, c'est l'objet du chapitre 8, ou bien à un niveau local, comme dans le chapitre 9.

8^{ième} CHAPITRE. Propositions de solutions liées au choix d'un critère de sélection global : la cartographie de *flux significatifs* ($F_{ij} > \alpha$)

« 2°) Ne pas représenter les mouvements au-dessous d'une certaine quantité, ce qui allège l'information d'un cortège d'impondérables généralement non significatifs. 3°) Ne pas représenter les mouvements au-dessous d'un certain pourcentage. Mais un très faible pourcentage peut cependant donner une quantité importante si la région est très peuplée ».

Jacques Bertin (1973).

« Puisqu'il n'est pas possible¹⁹⁹ de représenter tous les flux, l'approche consiste à les sélectionner, ce qui signifie qu'un certain nombre de flux – souvent la majorité – seront réputés moins saillants quel que soit le critère utilisé pour sélectionner les flux. »

Sandra Holland et David A. Plane (2001).

Introduction du chapitre

Ce huitième chapitre a pour objectif de présenter les solutions liées au choix d'un *critère de sélection global* à appliquer aux cellules de la matrice de flux, au préalable de leur cartographie. Cette procédure est rendue nécessaire lorsque la matrice est dense et que l'objectif n'est pas de représenter toute l'*information disponible* qu'elle contient, auquel cas, il est possible d'appliquer les solutions proposées dans le chapitre 5.

L'objectif de sélection des données de flux répond ici à deux objectifs. Le premier tient au besoin de réduire la complexité visuelle de la carte, correspondant à l'*effet de surcharge graphique*. Le second objectif est de sélectionner l'information de manière robuste afin que les cartes réalisées puissent être *comparées* avec d'autres soit parce qu'elles portent sur des catégories de flux relevant de la même zone d'étude, soit parce qu'elles concernent le même type de flux mesuré à différentes dates.

L'approche proposée est *globale*, ce qui signifie que le critère unique proposé agira à l'échelle de la matrice, simultanément sur l'ensemble des cellules et non par ligne(s) ou par colonne(s) ; les solutions liées à l'application d'un critère local (au niveau des lieux d'origine ou de ceux de destination) seront examinées dans le chapitre 9.

La matrice dont il est question contient indifféremment une information de flux (F_{ij}) ou de mouvement (M_{ij}), puisque nous avons vu dans le chapitre 6 que leur construction était identique sur le plan sémiologique, mais que leur interprétation différait.

Pour ces raisons, nous allons de nouveau nous placer dans une situation de corrélation quasi parfaite entre les trois mesures d'éloignement des lieux intervenant dans le processus cartographique à savoir (L_{ij}), (D_{ij}) et (E_{ij}). Nous allons également supposer que les déplacements réalisés dans l'espace se font à vitesse constante et en ligne droite, la projection cartographique et l'échelle de la carte respectent un rapport de proportionnalité strict entre distance géographique et distance cartographique ; la symbolisation des flux se fait à l'aide d'un trait droit faute de quoi la courbure

¹⁹⁹ Cette affirmation de S. Holland et D. Plane, énoncée il y a un peu plus d'une dizaine d'années, n'est plus tout à fait exacte aujourd'hui. Nous avons, en effet, montré (chapitre 5) qu'il était possible de cartographier l'ensemble de l'information disponible dans une matrice de flux, sans opérer de sélection ou de réduction des données a priori.

entraînerait des distorsions entre la distance cartographique et le déplacement spatio-temporel effectivement réalisé.

Nous rappelons, dans un premier temps, l'approche couramment réalisée de sélection globale des flux, par l'application d'un critère de sélection des flux défini *a priori*, de manière empirique. La carte décrit alors les *n plus gros flux* ($F_{ij} > n$). L'estimation du seuil étant empirique, la carte obtenue – tout en étant très efficace – ne conduit pas à un résultat qui soit robuste et surtout *comparable* avec d'autres. Pour cela, il convient de fixer le critère en utilisant des critères statistiques, ce que nous examinons dans la seconde section. Le recours à une approche statistique qui postule la normalité de la série (ce qui n'est pas le cas de celle des flux) va nous conduire à raisonner dans le cadre de ce chapitre sur des données log-normalisées $\log(F_{ij})$. Pour faciliter la démonstration, nous nous limiterons au cas de la matrice symétrique (F_{ij}^+).

L'objectif d'estimer un *critère unique de sélection* appliqué à l'ensemble de la matrice est un choix qui entraîne une double conséquence sur le plan théorique. Il conduit à accepter l'idée selon laquelle d'une part, certains lieux d'origine ou de destination des flux renseignés dans la matrice ne présenteront pas de lien (entrant ou sortant) sur la carte ; d'autre part qu'une partie des couples de lieux n'a pas d'intérêt pour l'analyse. Les flux intéressant les lieux dont il est question ne sont pas considérés comme *significatifs* dans le système de flux cartographié. Ils sont éliminés de la représentation, car ils ne répondent pas au *critère unique de sélection* que nous allons définir.

Il est important de noter que la part de flux *significatifs* sélectionnés peut fortement varier en fonction de l'existence d'une présélection réalisée dans le cadre d'un raisonnement selon la focale²⁰⁰ de la matrice.

L'approche globale vise à réduire la matrice aux seuls flux *significatif*, ceux qui répondent à un seuil (α) de *représentativité de l'information disponible* dans la matrice. Elle conduira nécessairement à l'élimination d'une partie des flux, par conséquent des lieux, ne répondant pas au critère. La question est donc de savoir **comment s'assurer que l'on cartographie l'information significative, en appliquant un critère unique de sélection à l'ensemble de la matrice, en assurant la comparabilité du résultat obtenu, puisque le raisonnement sous-jacent est identique à celui qui conduit à la réalisation de discrétisations statistiques** (l'application d'un seuil unique correspondant à une dichotomie). Existe-t-il une méthode permettant d'estimer la valeur du seuil de sélection des flux de manière que le résultat cartographique obtenu puisse être comparable entre deux matrices ?

Pour y répondre, nous allons suivre la suggestion de W. Tobler (1982) d'analyser la courbe de distribution des valeurs de flux dans la seconde section. Cela nous conduira à mobiliser trois outils distincts que sont l'histogramme de distribution des flux, leur transformation log ainsi que la courbe cumulative de concentration de ces valeurs.

Le recours à ce troisième graphique nous conduit à proposer, dans la troisième section, la mise en œuvre d'un double critère de sélection qui agit simultanément aux niveaux graphiques et analytiques, autrement dit sur la *densité* de figurés estimée par la part de liens représentés, ainsi que sur la *significativité* des valeurs représentées, par rapport à l'information totale. Nous verrons que la question de la sélection des flux, par un *critère unique*, conduit à réduire *la surcharge graphique* mais aussi la *couverture spatiale*. Elle intéresse le pourcentage cumulé de la quantité de figurés (L_{ij}) compte tenu du pourcentage cumulé d'information de flux pouvant être interprété soit en termes de flux (F_{ij}), soit en termes de mouvement (M_{ij}).

²⁰⁰ La sélection des flux selon une focale est une approche graphique et visuelle qui consiste à choisir la partie de la matrice à analyser (les lieux, les lieux, certains lieux...) et non de raisonner selon les valeurs de ses cellules.

8.1 Le critère ($F_{ij} \geq \alpha$) de la sélection directe des flux à cartographier

Une pratique courante de sélection des flux à représenter consiste à ne retenir que les n plus gros flux. Cette procédure étant largement répandue, nous n'allons pas nous y attarder et la présentons succinctement, avant de généraliser le propos en considérant le rang des valeurs : $\text{rang}(F_{ij}) \leq n \leftrightarrow (F_{ij} \geq \alpha)$.

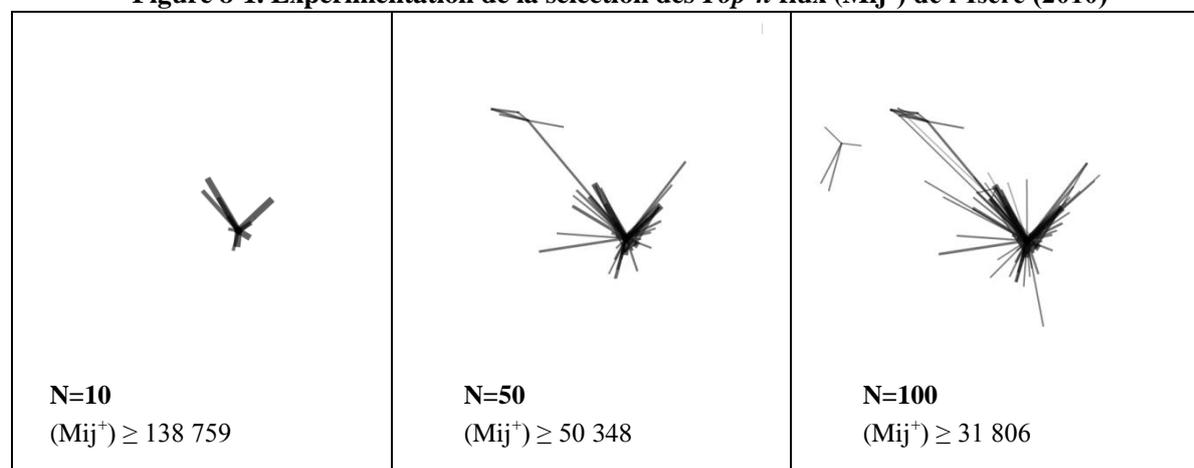
8.1.1 La cartographie du Top- n des flux

Cette procédure nécessite un tri préalable des valeurs en ordre décroissant, puis l'application d'un seuil généralement empirique ($\alpha = \text{Top-}n$) au-delà duquel les données seront représentées. A noter que dans le cas où les flux sont décrits du point de vue des liens et des lieux, il n'est pas nécessaire de sélectionner les lieux : tous seront représentés.

Cette pratique très commune est souvent mise en œuvre pour cartographier les flux qui s'expriment à l'échelle mondiale. Le critère utilisé est alors celui des 50 (*Top-50*) voire des 100 (*Top-100*) plus gros flux, supposés concerner les lieux qui présentent une attractivité à l'échelle mondiale. Il est d'ailleurs souvent appliqué pour sélectionner les flux, indépendamment du fait qu'ils soient observés du point de vue des liens ou des lieux. N. Cattani (2004) avait ainsi retenu le top-100 (les 100 premiers flux) des flux internationaux de passagers par voie aérienne (voir chapitre 7) et C. Grataloup (2011), les 50 plus gros flux mondiaux de marchandises (voir chapitre 6).

Les cartes de flux ainsi réalisées décrivent une part réduite de l'information et cela, quelle que soit l'échelle à laquelle la procédure est appliquée. En opérant ainsi, les cartes éludent de la représentation un certain nombre de lieux d'autant plus important que la valeur du seuil est élevée. Pour l'illustrer, nous avons réalisé une courte expérimentation sur la matrice (F_{ij}^+) des navettes de l'Isère. La Figure 8-1 présente l'application des trois critères de sélection suivants des n valeurs de flux les plus importantes : $n > 10$; $n > 50$ et $n > 100$.

Figure 8-1. Expérimentation de la sélection des Top- n flux (M_{ij}^+) de l'Isère (2010)



Source : fichier de détail, MOBPRO, INSEE, 2010. 54 973 273

Les images de la Figure 8-1, comparées avec celle de la Figure 5-7 (page 225) qui décrivait toute l'information disponible dans la matrice (F_{ij}^+) de l'Isère, mettent en évidence la réduction drastique de l'information liée au critère de sélection de type *Top- n* . Elles illustrent les limites de cette forme de seuillage qui apparaît de fait peu pertinente à l'échelle locale. En effet, la sélection des dix (*Top-10*) voire des 50 (*Top-50*) plus grosses valeurs de flux de l'Isère ne décrit que quelques navettes locales structurées autour de la préfecture du département, Grenoble. Elle ne met pas en évidence les principaux motifs des flux de cette matrice préalablement identifiés (voir Figure 5-10, page 229).

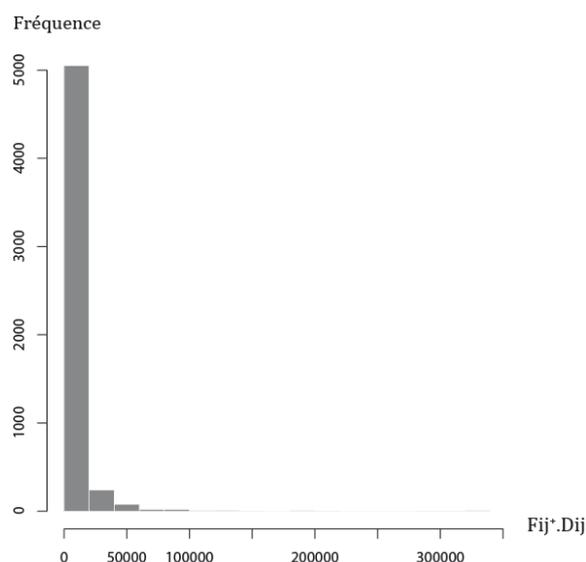
La pertinence du critère de sélection de type *Top-n* n'est pas avérée, la valeur du critère devant être réalisée en lien avec la résolution des données et la thématique du flux, afin de produire à peu de frais, des cartes qui soient efficaces. Indépendamment de cela, le principal problème du *top-n* réside dans la production de cartes qui ne sont ni reproductibles, ni comparables entre elles, le critère n'étant pas fondé sur l'examen de la distribution des valeurs de flux. W. Tobler (1982) a, en effet, suggéré de visualiser la distribution des valeurs de flux au préalable de leur exploitation, de la même façon que l'on analyse des données statistiques avant de les discrétiser avant leur cartographie. Aussi d'adopter à minima « *une stratégie optimale [qui] consiste à retirer tous les flux dont la valeur est inférieure au flux moyen total* » (Tobler, 1982).

8.1.2 Le critère ($F_{ij} \geq a$) : de l'examen de la forme de la distribution des valeurs de flux à l'étude de leur concentration

L'examen de la forme de la distribution de données statistiques est une opération courante en cartographie statistique. Elle est étrangement peu mise en œuvre dans l'analyse des flux, alors qu'il s'agit d'une étape pourtant essentielle, qui détermine les éventuels traitements à réaliser sur les données. Elle conditionne aussi, dans bien des cas, les choix réalisés en matière de cartographie, en particulier ceux qui portent sur la discrétisation des valeurs et sur l'application des principes de sémiologie graphique.

L'examen de la distribution des flux fait partie des méthodes graphiques qui ne présentent aucune difficulté particulière quant à leur mise en œuvre. Elle consiste à réaliser une courbe qui décrit la forme de la série de données analysées. Nous présentons successivement les trois principaux outils qui peuvent être mobilisés, à savoir l'histogramme de distribution des valeurs de flux, celui des fréquences cumulées ainsi que la courbe de concentration des valeurs de flux. Le tracé d'une courbe de distribution a pour objectif de visualiser la manière dont les valeurs sont réparties. Elle ne présente aucun intérêt analytique dans le cadre d'analyse des flux, puisque la figure n'est pas exploitable en l'état. A titre d'exemple, la Figure 8-2 présente l'histogramme obtenu des valeurs de (M_{ij}^+) de l'Isère.

Figure 8-2. L'histogramme des valeurs de flux (M_{ij}^+) de l'Isère

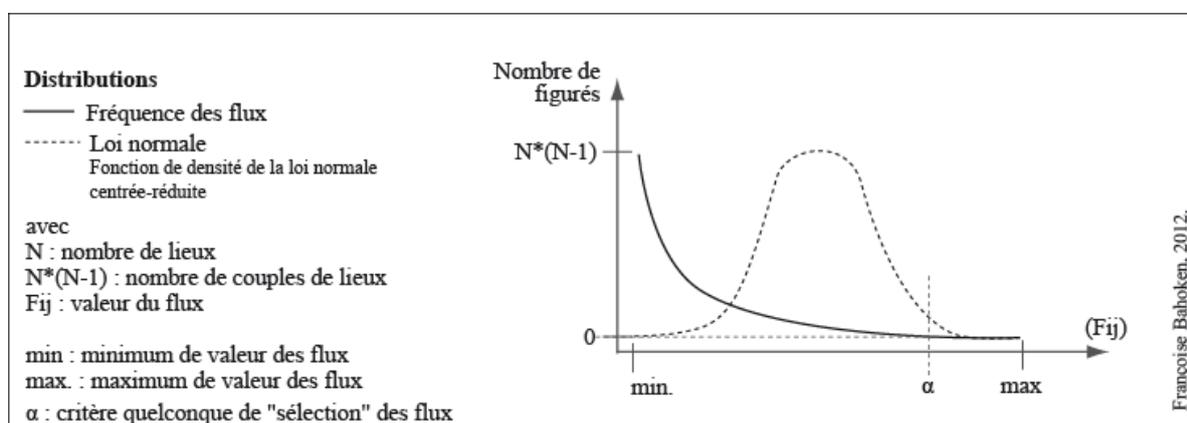


Cette figure est caractéristique de la très forte dissymétrie à gauche des données de flux. Elle traduit l'existence de nombreux petit flux, c'est-à-dire de liens qui présentent des valeurs faibles, et de peu de liens qui présentent des valeurs de flux élevées. Cette courbe présente une forme dite puissance ou

Pareto, éloignée de la forme gaussienne classique caractéristique de la Loi normale²⁰¹ constituant le cadre d'analyse habituel de la statistique descriptive. Il n'est donc pas possible d'appliquer de traitement statistique qui postule la normalité sur une telle distribution.

La Figure 8-3 présente la distribution théorique des valeurs de flux à la loi normale.

Figure 8-3. La courbe de distribution théorique des valeurs de flux



Cette figure décrit la relation entre la quantité de liens, placée sur l'axe des ordonnées (Y) et les valeurs de flux, placées sur l'axe des abscisses (X).

L'allure dissymétrique de la série des flux est liée à la manifestation de différents effets spatiaux qu'il n'est pas envisageable de présenter dans les détails. Ces effets se répercutent sur la carte, en générant cette complexité visuelle qui caractérise l'*effet-spaghetti*. C'est la fréquence trop importante de faibles valeurs de flux — dont nous verrons qu'ils n'apportent pas d'information notable — qui caractérise l'illisibilité de la figure, et justifie le recours à un critère de sélection.

Ce constat n'étant pas spécifique à la matrice des flux de l'Isère, certains auteurs ont fait l'hypothèse que **la distribution des valeurs de flux pouvait se rapprocher d'une courbe gaussienne, sous réserve que deux conditions soient simultanément remplies :**

- l'exclusion des valeurs nulles ($F_{ij} < 0$) de l'analyse, afin d'éliminer le pic de valeurs situé sur la partie gauche de l'histogramme (voir Figure 8-2) ;
- une transformation de ces valeurs de flux (F_{ij}) en $\log(F_{ij})$.

« Faire une transformation, c'est utiliser une fonction de la variable au lieu de la variable elle-même. C'est, en fait, simplement changer d'échelle et transférer les mesures effectuées sur une échelle originelle sur une autre échelle. Le résultat est donc, d'après Taylor (1977) de convertir sur l'échelle, les intervalles entre les observations ». (Cauvin et al., 1979 : 51-52).

Le choix du type de transformation à appliquer aux valeurs de flux pour les normaliser est lié à l'importance de la dissymétrie de la série, comme l'indique (Poitevin, 1999 : 117-118). Ainsi, si la distribution des flux est :

- peu dissymétrique, il convient d'appliquer la fonction racine carrée : $F_{ij} \leftarrow (F_{ij})^{1/2}$;
- dissymétrique, c'est plutôt une fonction racine cubique $F_{ij}^* \leftarrow (F_{ij})^{1/3}$ ou une fonction logarithmique $F_{ij} \leftarrow \log(F_{ij})$;
- très dissymétrique, la fonction inverse $F_{ij} \leftarrow 1 / (F_{ij})$ est la plus adéquate, mais elle inverse la position des valeurs fortes et faibles.

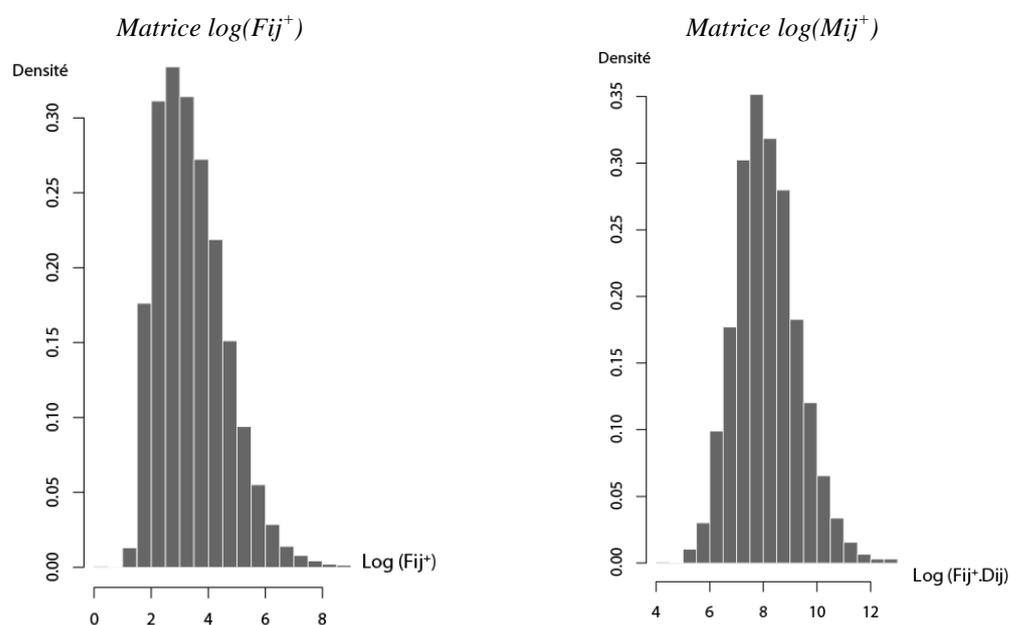
²⁰¹ Ce recours à la loi normale est très aisé pour qui n'est pas statisticien, car il permet l'accès à une variété de traitements qui ne sauraient lui être accessibles autrement.

Les transformations des valeurs de flux les plus utilisées sont la fonction racine cubique ou la fonction logarithmique. Nous faisons le choix de la fonction logarithmique qui s'avère être la plus courante, ce qui nous conduit à examiner l'histogramme des valeurs de flux transformées $\log(F_{ij})$.

La distribution des valeurs de $\log(F_{ij})$ est globalement log-normale, ce qui signifie que son histogramme de distribution suit à peu près la forme gaussienne.

La Figure 8-4 présente les histogrammes des valeurs log normales des matrices $\log(F_{ij}^+)$ et $\log(M_{ij}^+)$ de l'Isère, correspondant respectivement aux quantités de flux et de mouvements. Elle considère une mesure de *distance géographique* (D_{ij}) exprimée en kilomètres, prise dans la même situation idéale de corrélation parfaite ou quasi parfaite entre les trois mesures d'éloignement des lieux envisageables que sont : la *distance-temps* exprimée en minutes ; la *distance géographique* en kilomètres et la *distance cartographique* en millimètres.

Figure 8-4. Histogrammes des valeurs transformées des matrices (F_{ij}^+) et (M_{ij}^+) de l'Isère



Les deux distributions présentées sur la Figure 8-4 sont désormais uni-modales et globalement bien réparties autour de la valeur moyenne. On observe toutefois que l'ajustement de la distribution (M_{ij}^+) à une loi log-normale semble un peu meilleur que celui de (F_{ij}^+) , car la symétrie est plus nette par rapport à la moyenne.

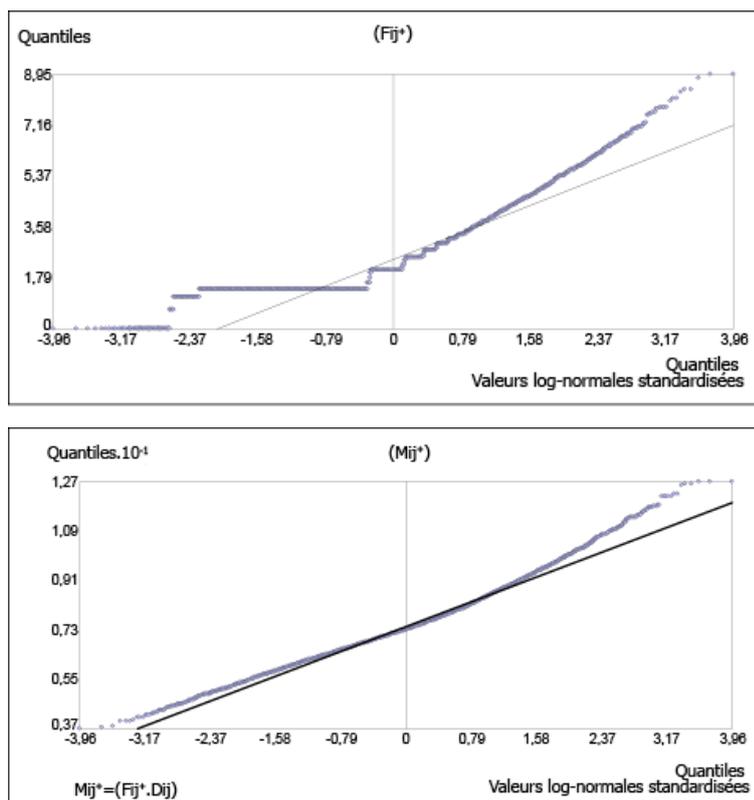
L'avantage théorique du recours au $\log(F_{ij})$ est double et n'est pas spécifique aux flux. D'une part, les cartes réalisées pourront être comparées avec d'autres, si la méthode est appliquée sur des matrices complexes : des jeux de données qui portent sur la même zone d'étude qui sont disponibles à plusieurs dates ou qui sont décomposables en catégories, de même que ceux qui concernent des zones d'études différentes. D'autre part, le choix du critère de sélection pourra être fondé sur une discrétisation de la série qui sera plus robuste, pour les mêmes raisons qu'une transformation effectuée sur des données scalaires dissymétriques — au préalable d'une discrétisation par standardisation — produit une carte plus facile à comparer qu'un seuillage dit naturel. Nous avons déjà vu dans le chapitre 7 l'intérêt du recours à une discrétisation pour segmenter la matrice dans l'objectif de réaliser une collection de cartes.

Le moindre ajustement de la matrice $\log(F_{ij}^+)$ n'est pas spécifique aux données de l'Isère, mais général au cas de flux. Il résulte du fait que la distribution des flux ne fait que tendre vers une forme log-normale, la valeur du coefficient d'asymétrie obtenu après transformation n'étant pas nul. Pour la

matrice $\log(M_{ij}^+)$ de l'Isère qui apparaît la mieux ajustée, ce coefficient est de 0,55, alors qu'il était de 14,7 pour la matrice (F_{ij}^+) . Ce coefficient non nul suggère une déviation du $\log(F_{ij})$ par rapport à la loi normale préjudiciable à la détermination du seuil. La représentation de cette déviation est illustrée sur la Figure 8-5, sous la forme d'un diagramme Quantile-Quantile (Q-Q).

Le diagramme Q-Q de la courbe des $\log(M_{ij}^+)$ met en évidence deux problèmes qui traduisent le fait que l'ajustement à la loi normale n'est pas effectif sur l'ensemble de la série. On observe, en effet, sur la Figure 8-5 une déviation de la courbe d'ajustement des valeurs à la loi normale sur les bords de la série (au-delà de -1,58 et de +1,58).

Figure 8-5. Diagramme Q-Q des valeurs de $\log(F_{ij}^+)$ et $\log(M_{ij}^+)$ de l'Isère



Le second problème réside dans la variation de cet écart en fonction du type de matrice, pour une même transformation. L'ajustement à la courbe log-normale est encore moins bon pour les matrices de type (F_{ij}) et (F_{ij}^+) alors qu'il tend vers la normale pour $\log(M_{ij})$ et $\log(M_{ij}^+)$.

Ce constat d'un ajustement imparfait par la fonction logarithme étant avéré, les résultats cartographiques obtenus suite à l'application d'un critère de sélection sur ces séries – en utilisant des méthodes qui postulent leur normalité – ne sauraient être considérés comme totalement satisfaisants. Deux possibilités d'investigation sont dès lors envisageables pour améliorer les résultats cartographiques de (F_{ij}) .

Une première direction consiste à utiliser d'autres fonctions d'ajustement qui seraient mieux adaptées à la forme dissymétrique de la distribution des valeurs de flux. L'investigation serait alors menée à un niveau statistique qui dépasse le cadre de notre analyse, c'est pourquoi elle ne sera pas engagée. La seconde voie consiste à examiner la courbe de fréquences cumulées des valeurs de flux (F_{ij}) , au lieu de celle des fréquences simples. Nous proposons d'examiner cette seconde possibilité, dans la section suivante, en mobilisant pour cela un troisième outil graphique : la courbe de concentration (ou courbe de Lorenz) des fréquences cumulées des valeurs flux.

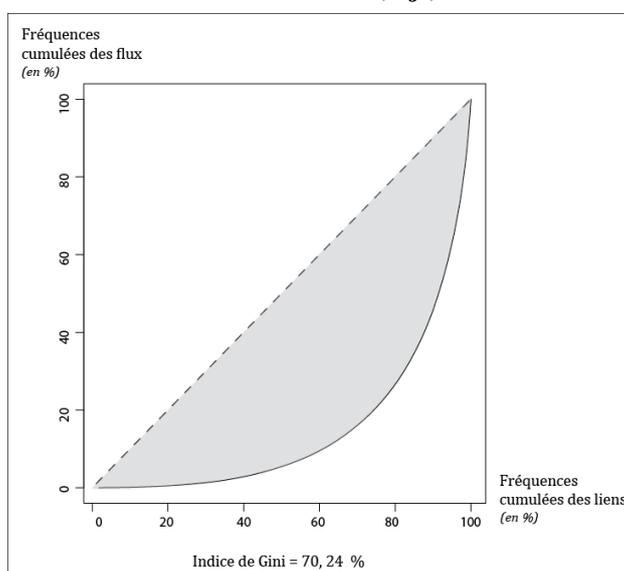
8.2 Le critère de concentration (Lorenz) des fréquences de valeurs de flux : (F_{ij} vs L_{ij})

La courbe de Lorenz permet de représenter sur une même figure la répartition des fréquences cumulées des individus (les liens) et de leurs valeurs, en pourcentage de l'information totale qu'ils représentent respectivement. Elle est traditionnellement utilisée pour évaluer l'inégalité d'une distribution statistique qui, lorsqu'elle est appliquée au cas des flux, correspond à l'existence d'une dissymétrie entre la part en pourcentage de liens (L_{ij}) et la part de l'information totale de flux (F_{ij}) ou de mouvement (M_{ij}) qu'ils représentent. Elle permet de faire le lien entre les aspects analytiques et graphiques de la carte de flux : entre la part d'information qui est représentée sur la carte et la part de liens équivalente.

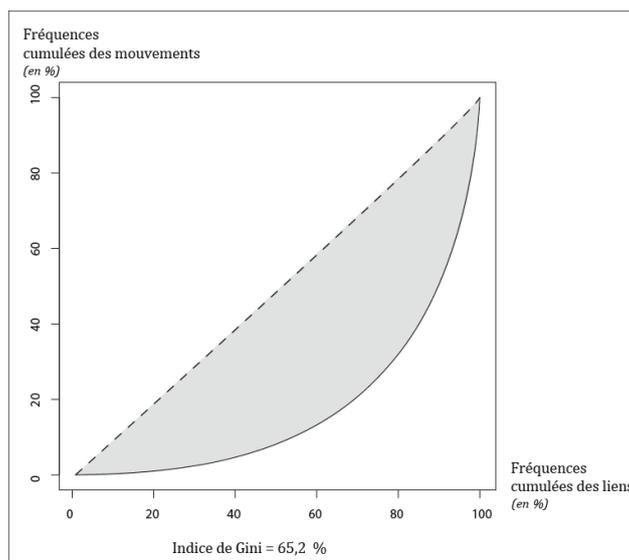
La Figure 8-6 présente les courbes de concentration statistique des valeurs du volume de flux (F_{ij}^+) et de mouvement (M_{ij}^+). Elles sont exprimées en fonction du % cumulé de liens.

Figure 8-6. Les courbes de concentration de flux (F_{ij}^+) et de mouvements (M_{ij}^+) de l'Isère

Matrice (F_{ij}^+)



Matrice (M_{ij}^+)



La courbe de Lorenz décrit l'inégalité éventuelle de la fréquence des valeurs de flux, qui est par ailleurs mesurée par l'indice de concentration de Gini.

Cet indice de Gini correspond, sur le plan graphique, à la mesure de la partie du graphique portée en gris sur la Figure 8-6, ce qui correspond au degré de concentration de la matrice représentée : 70,4 pour (F_{ij}^+) et 65,2% pour (M_{ij}^+) . L'indice varie entre 0 et 1, lorsqu'un seul individu détient toute l'information : dans ce cas, l'égalité de la fréquence des valeurs est parfaite. Plus la courbe de Lorenz s'éloigne de la diagonale et plus l'indice de Gini est élevé, et la distribution concentrée sur quelques individus qui contribuent fortement à l'information. **Dans le cadre d'analyse des flux, l'indice de Gini permet donc de mesurer la concentration des valeurs fortes de la matrice sur quelques individus, ce qui revient à appliquer simultanément les deux logiques (de flux et de liens) caractéristiques du critère unique de sélection. La concentration étant interprétée, sur le plan analytique, en termes de concentration des valeurs de flux sur une quantité faible de liens. Ainsi, plus l'indice de Gini d'une matrice de flux est élevé et plus on peut espérer visualiser un maximum d'information avec un minimum de liens.**

Pour faciliter l'examen de cette courbe, il serait d'ailleurs intéressant de pouvoir disposer d'une interface interactive qui aiderait l'utilisateur à la détermination de ce critère de sélection, de manière visuelle (directement sur le graphique) ou *a priori*, en indiquant la valeur du critère unique de sélection. Une proposition pour la réalisation d'une telle interface est présentée dans la section 8.4.

La Figure 8-7 présente les résultats qui pourraient être obtenus à la suite de l'étape de sélection globale des valeurs de flux à représenter, sur la matrice (M_{ij}^+) de l'Isère.

En dessous de chacune des images est portée la valeur (α) du critère de sélection des flux : $(M_{ij}^+) \geq \alpha$. Sur les seconde et troisième lignes figurent la part de liens et la part de flux, en % du total dont il est question. Seuls les liens qui présentent une valeur supérieure ou égale au critère (α) sont représentés, ce qui a pour conséquence que certains lieux sont éliminés de l'analyse puisqu'ils ne présentent plus de liens.

Ainsi, la première image de la Figure 8-7 (en haut à gauche) correspond à la représentation de 8,5% de l'information totale (qui correspondent aux valeurs de flux supérieures ou égales à 100 $(M_{ij}^+) \geq 100$) par 0,07% des liens. Seuls les liens qui présentent une valeur supérieure ou égale au critère $\alpha=138\ 759$ sont représentés.

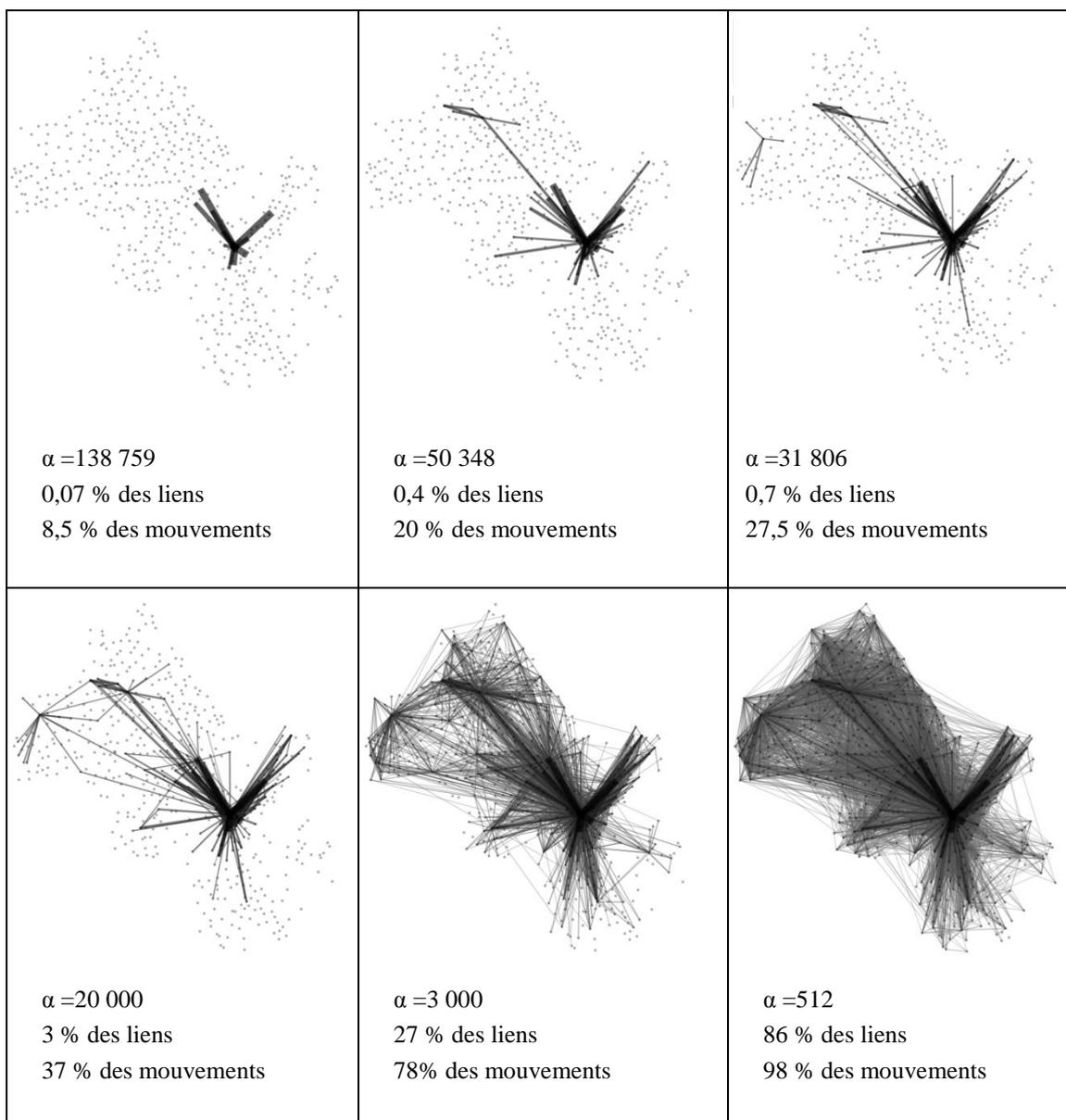
Sur le plan thématique, ces images illustrent la masse des petits flux qui parcourent de longues distances, ce qui est normal puisque la matrice (M_{ij}^+) est à interpréter ici selon une *logique de mouvement*. La polarisation de ces mobilités quotidiennes de l'Isère autour des pôles d'emplois demeure. Si la première image ne décrit que l'aire d'influence du pôle grenoblois, l'exploration interactive met rapidement en évidence les flux inter pôles qui font fi de la topographie contrainte des lieux.

Les volumes bilatéraux de flux les plus importants concernent des actifs qui vont travailler dans les communes situées au cœur de l'agglomération, sur le replat de la cuvette grenobloise.

Cette figure illustre l'intérêt du recours à une courbe de concentration, pour déterminer visuellement la valeur du critère de sélection, compte tenu du résultat cartographique obtenu. En effet, lorsque les valeurs de flux sont triées par ordre décroissant, la courbe permet de sélectionner un pourcentage de liens par rapport à un pourcentage de flux par classe qui correspond à l'estimation du seuil, nécessairement interprétée de manière visuelle au regard de la carte de flux obtenue.

Ce constat nous conduit à la conclusion suivante : le raisonnement qui préside à la détermination du critère (α) de sélection globale des flux à représenter résulte d'un arbitrage entre la part de liens et la part d'information de flux représentée. Cela a pour conséquence que **le critère de sélection globale n'est pas exactement unique. Il correspond, en réalité, à un double critère qui est défini a priori, de manière unique et inconsciemment, en suivant deux logiques de raisonnements distinctes.**

Figure 8-7. Exploration « interactive » de l'effet cartographique de l'application du double critère de sélection sur la matrice (M_{ij}^+)



Source : fichier détail MOBPRO, INSEE, 2010.

Ce double critère est cependant toujours objectivé puisque nous pouvons démontrer qu'il correspond à un raisonnement fondé soit sur la sélection des flux (ou des mouvements), soit sur celle des liens. Ces deux logiques agissent de la même manière sur le même critère : elles sont mises en œuvre dans le même objectif de cartographie de flux qui sont *représentatifs* de l'information disponible, de ceux qui sont considérés comme *significatifs*.

La question qui se pose à ce stade est alors celle de savoir comment estimer la valeur du critère (α) : en suivant une *logique de liens* ou bien de *flux / de mouvements* ? Les résultats obtenus sont-ils invariants quelle que soit la logique du raisonnement adoptée ? Quel est l'intervalle plausible de valeurs prises par (α) ?

8.3 L'estimation du double critère de sélection ($F_{ij}, L_{ij} > \alpha$) de la part (en %) de l'information de flux en fonction de la part de liens représentée et inversement

L'application d'un *critère global de sélection unique* des valeurs de flux à représenter consiste, en réalité, à réaliser une double sélection des flux combinant les logiques de liens et flux : au regard de la part de liens (L_{ij}) par rapport à la part d'information rapportée au total de flux (F_{ij}) ou bien de mouvement (M_{ij}). Ainsi formulé, **l'enjeu de la détermination de ce critère, pour les matrices de flux denses qui présentent un indice de Gini élevé, réside donc dans la cartographie optimale du maximum d'information de (F_{ij}) ou de (M_{ij}) par un minimum de liens (L_{ij}), afin de garantir clarté, efficacité visuelle et analytique de la figure.**

L'efficacité définie sur le plan visuel est d'ordre graphique, puisqu'elle consiste à réduire *l'effet spaghetti*, en particulier la *surcharge graphique*, en éliminant les liens considérés comme peu voire *non significatifs* pour l'analyse. Sur le plan analytique, l'efficacité de la carte réside dans sa capacité à donner à voir les principaux motifs représentés, ceux qui font sens dans la thématique dont il est question, selon qu'il s'agit de flux (migratoires, financiers, commerciaux, ...) ou de mouvements. En effet, la cartographie issue de la sélection des 100 premières quantités de flux ne sera pas identique à celui des 100 premières quantités de mouvement et cela, pour des raisons qui tiennent à l'interprétation de la carte – comme nous l'avons vu dans le chapitre 6 – mais aussi à la manière dont la sélection des données à été réalisée.

La question du seuillage significatif d'une série statistique est un problème ancien et bien connu dans le champ du traitement de données. Il se pose avec acuité pour les flux depuis plusieurs années, sans toutefois avoir été résolu (Becker, Eick, Miller, Wilks, 1990 : 93) cités par (Dorling, 1991 : 112). Deux raisons peuvent être avancées pour expliquer l'absence de critère de sélection unique des flux. La première tient certainement à la mise en œuvre d'approches statistiques qui sont inadaptées. Les statistiques paramétriques classiques postulent la log-normalité de la distribution de flux, pour déterminer la valeur du critère de sélection, ce qui n'est pas tout à fait le cas, et ce constat est valable sur la plupart des matrices de flux, au-delà du cas de l'Isère. Le fait que l'histogramme de distribution du $\log(F_{ij})$ ne soit pas parfaitement ajusté à la loi normale (voir Figure 8-5, page 331) ouvre la voie vers la recherche théorique de solutions et de leur application au cadre de l'analyse cartographique des flux, mais qui sont actuellement hors de notre portée²⁰². La seconde raison tient, de notre point de vue, au caractère relatif du critère de sélection des flux. Ce *double critère unique* ne correspondrait pas à une valeur exacte, invariante, quelle que soit la matrice, mais s'inscrirait dans un intervalle de valeurs non nécessairement borné. La recherche d'une valeur optimale et unique de seuillage serait donc vaine.

Toutefois, nous avons pu observer sur différentes matrices que la valeur du critère de sélection des flux variait peu dans un certain intervalle de valeurs (nous y reviendrons dans la section 8.3.2). Nous en concluons que le seuil ne correspond pas à une valeur exacte et que sa recherche est un vrai faux problème. La question de l'estimation du critère de sélection est plutôt soit celle de l'estimation d'un intervalle²⁰³ de valeurs les plus fortes, c'est-à-dire concentrées sur l'histogramme, soit inversement, celle de la dispersion des liens ; tout dépend de la logique du raisonnement adopté. L'objectif étant bien de produire, quelle que soit la logique, une carte de flux qui soit efficace sur les plans graphiques et analytiques.

²⁰² Nous pensons que les statisticiens disposent certainement de fonctions de transformation qui permettent de normaliser parfaitement ces distributions de flux (F_{ij}), autrement que par la fonction logarithme.

²⁰³ L'exploration de l'effet cartographique de l'application du double critère de sélection (dont quelques résultats sont présentés sur la Figure 8-7) nous a conduits à la conclusion selon laquelle la carte de flux ne variait pas dans un certain intervalle de valeurs de flux ou de liens. C'est aussi la raison pour laquelle nous avons suggéré l'adoption d'une logique floue, une première analyse étant présentée dans (Bahoken, 2012, 2014).

Il y a sans doute là l'un des éléments d'explications du fait que les cartographes ne mentionnent pas, dans l'ensemble, comment ils en arrivent aux valeurs de seuils des cartes de flux qu'ils produisent. En effet, que signifie la phrase suivante : « *Je représente 80% des flux* » ? Si nous ne doutons pas du caractère significatif des 80% de flux sélectionnés par l'auteur d'une carte, nous souhaitons toutefois souligner l'aspect « boîte noire » du raisonnement sous-jacent à l'estimation de ce critère. Le problème qui se pose à ce stade n'étant pas la valeur de ce critère – puisque nous faisons l'hypothèse qu'il est défini sur un intervalle –, mais la famille de solutions à apporter pour son estimation, dont nous pensons qu'elle est d'abord d'ordre méthodologique. En faisant l'économie de l'explicitation du critère de sélection utilisé, le caractère reproductif et l'éventuelle comparabilité des cartes de flux réalisées ne sont plus assurés. Et pourtant, le choix qui conduit à la représentation des 80% de flux les plus *significatifs* n'est pas dénué de fondements.

Le critère de sélection des flux à cartographier, souvent défini *a priori* et inconsciemment est, de notre point de vue, toujours objectivable. Pour preuve, il est généralement exprimé sous deux formes sur les cartes : il correspond, par exemple, à la représentation de ces *80% des flux* ou bien à celle des *2% des flux* les plus importants, les deux expressions renvoyant à des réalités distinctes. Ces deux types de précisions apportées quant à la part de flux représentée sur la carte correspondent en fait à deux logiques de raisonnements que nous avons pu identifier. La première logique s'inscrit dans un raisonnement mené sur les valeurs des flux, la carte représente alors 80% du total des flux : elle est réalisée en réponse à un objectif de *significativité* de l'information représentée. Dans la seconde logique, la carte représente 2% des liens disponibles dans la matrice, en l'occurrence 2% des liens les plus forts : elle répond à un *objectif de densité* de la matrice.

Nous présentons ces deux logiques de manière distincte dans les deux sous-sections suivantes, sachant qu'elles sont en pratique liées puisqu'elles participent à la définition du même critère unique de sélection. Le choix d'une présentation distincte n'ayant d'intérêt que pour les besoins de la validation des deux hypothèses suivantes : **le double critère unique de sélection s'inscrit dans un intervalle de valeurs de flux à l'intérieur duquel la figure varie peu ; le résultat cartographique ne varie pas en fonction de la logique du raisonnement adoptée pour le déterminer.**

8.3.1 L'objectif de significativité des valeurs ($\alpha \approx 80\%$)

La détermination du critère de sélection des flux, fondée sur un objectif de *significativité* des valeurs, vise à représenter la part maximum d'information, sur une carte lisible. Nous avons pris l'exemple précédemment de 80% de l'information totale, ce qui signifie que les 20% restants sont considérés comme résiduels.

La représentation de 80% de l'information considérée comme *significative* pour l'analyse, ou *représentative* (en %) de l'ensemble revient à appliquer la Loi de répartition des 80/20 spécifique aux distributions statistiques de forme Paréto. Le transfert de ce principe à la cartographie d'une matrice de flux conduit en fait à l'hypothèse plausible selon laquelle **une faible part de liens contribue à décrire une part importante de l'information de flux**. C'est probablement la raison pour laquelle différents auteurs (Tobler, 1982 ; Dorling, 1991 ; Bopda et al, 2000) s'y réfèrent pour sélectionner la part de flux cartographiée (voir Encart 8-1).

Encart 8-1. Exemples de mise en œuvre de la règle des 80/20

W. Tobler (1982) considère précisément la règle des 85-25% des valeurs de flux à représenter, arguant du principe que 85% du total des flux sont bien décrits par 25% des figurés. D. Dorling (1991) propose de ne représenter que de l'ordre de 80% de l'information qu'il considère significative, afin que « [...] *plus des trois-quarts des flux soient représentés sur la carte, les plus typiques étant ceux des banlieusards. Cette technique est particulièrement utile lorsque tous les flux inter districts de la Grande-Bretagne sont représentés, car il y a trop peu d'espace [sur la carte] pour utiliser efficacement la technique de l'ordonnancement visuel des figurés, les largeurs des lignes ne pourront pas être discriminées autrement* » (Dorling, 1991 : 111). A. Bopda et al. (2000), après avoir présélectionné les flux dits majeurs (sur lesquels nous reviendrons dans le Chapitre 9), représentent les « [...] *échanges migratoires [...] qui totalisent dans chaque cas plus de la moitié des échanges migratoires observés* » (Bopda, Grasland, Poulain, 2000 : 113), ce qui correspond environ à 60% de l'information disponible représentée par 5% des flux. N. Cattan suit une démarche similaire, en ne représentant que le top-100 des flux majeurs de l'OACI. En nous plaçant dans des conditions similaires (Voir Chapitre 7, section 7.4) à celles de l'auteur, la cartographie de ces 100 premiers flux revient à ne représenter que 15% de l'information totale par 1% des liens, soit les « [...] *principaux flux de passagers [qui] qui s'effectuent entre les trois ensembles territoriaux que sont les Etats-Unis, l'Europe occidentale et les trois dragons du sud-est asiatique, [...]* » (Cattan, 2004, *op.cit.*).

La Figure 8-8 illustre l'application de cette règle des 80/20 à deux matrices de flux de l'Isère : (Fij⁺) et (Mij⁺). Sur la partie haute de la figure, les flux sont sélectionnés en fonction d'un objectif de représentativité de 80% de l'information totale, tandis que sur la partie basse, l'objectif est fixé à 20% des liens. L'objectif est de montrer qu'il n'y a pas de différence notable, au niveau du rendu cartographique, selon que la règle des 80/20 est mise en œuvre du point de vue des flux ou des liens, les cartes ne variant pas sensiblement. Ce principe fonctionne, en effet, quel que soit le degré de concentration de la matrice.

Ce résultat valide aussi le fait que le critère de sélection s'inscrit dans un intervalle de valeurs à l'intérieur desquelles l'image varie peu. En effet, pour la matrice (Fij⁺), par exemple, le résultat est similaire selon que ($\alpha=60$) ou ($\alpha=81$) qui correspondent respectivement à la sélection des 80% de flux ou des 20% de liens

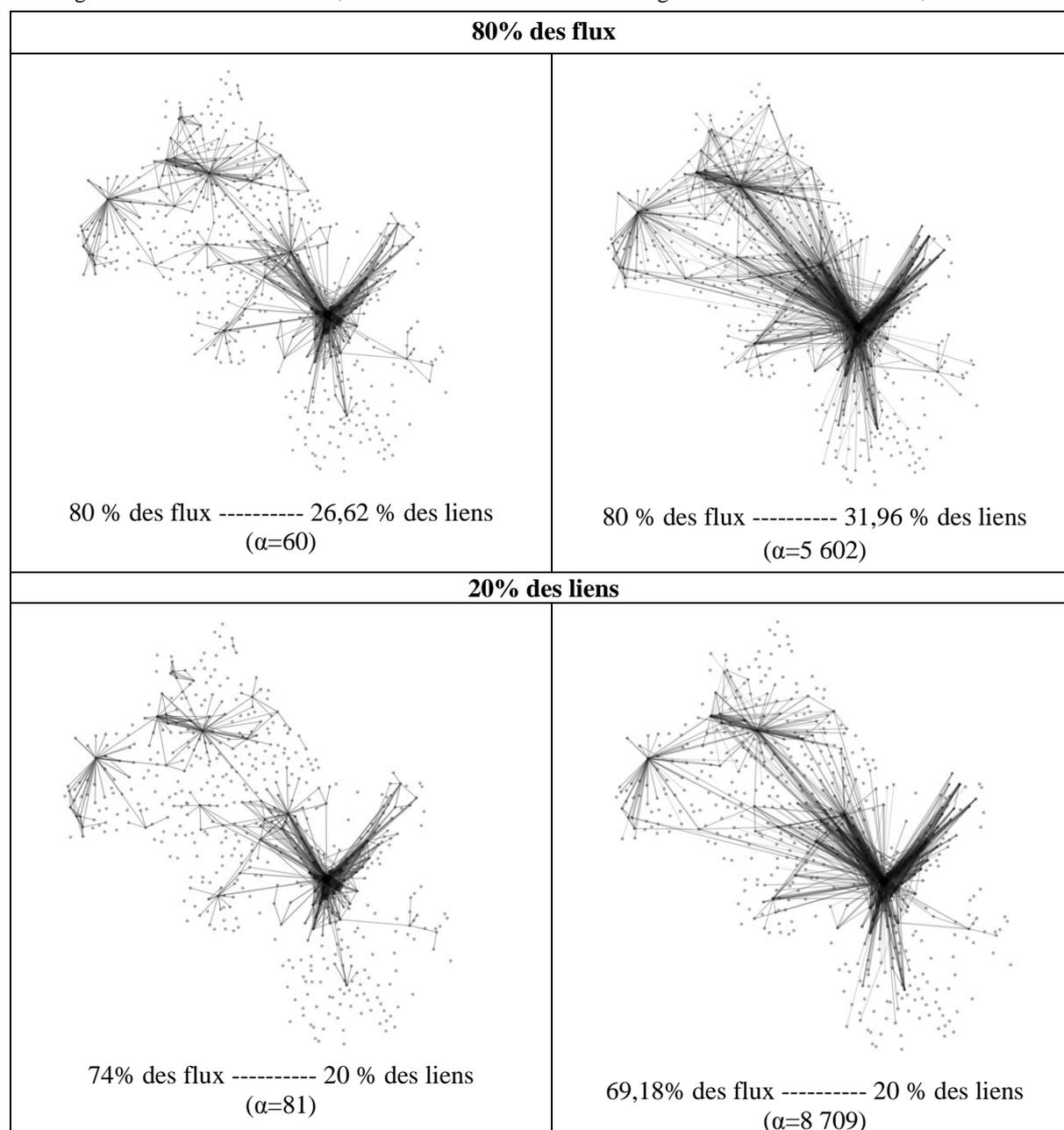
Figure 8-8. Cartographie des 80% de flux significatifs des matrices (F_{ij}^+) et (M_{ij}^+) de l'Isère

Matrice (F_{ij}^+)

Indice global de concentration : 70,24

Matrice (M_{ij}^+)

Indice global de concentration : 65,2



Le second type de raisonnement qui peut être appliqué pour déterminer le critère de sélection s'appuie sur un *objectif de densité* de la matrice. Il aboutit également à la représentation des quelques flux que l'on considère comme les plus *significatifs*. Le raisonnement examiné dans la sous-section suivante est donc l'inverse de celui que nous venons de présenter, tout en lui étant lié.

8.3.2 L'objectif de densité des valeurs : (2 – 5) %

L'idée qui consiste à déterminer la valeur du critère de sélection des flux en fonction d'un *objectif de densité* de la matrice est fondée sur un raisonnement graphique. Elle s'appuie sur le nombre de cases renseignées, donc sur le nombre de liens disponible pour l'analyse par rapport au nombre de liens

théorique. Cette notion de densité de la matrice est, en effet, empruntée à l'analyse des réseaux sociaux, issue de la théorie des graphes.

En effet, la densité est l'un des indicateurs classiques permettant de caractériser un réseau de relations. L'indice (γ) gamma, souvent utilisé pour les réseaux spatialisés, exprime, par exemple, le rapport entre le nombre de liens – qui seront représentés – sur le nombre de liens possibles, soit une densité de remplissage de la matrice. Cet indice varie entre 0 et 1 lorsque le réseau est complet, c'est-à-dire en présence de l'ensemble des liaisons théoriquement disponibles, soit $N*(N-1)$. Plusieurs formulations de la densité d'un réseau spatial, éventuellement valué, existent dans la littérature. Toutes peuvent être interprétées comme une forme de sélection des valeurs qui s'applique, par extension, aux matrices de flux.

Pour les flux, le critère généralement admis est de l'ordre de 0,5% de la densité de la matrice, lorsqu'elle est complète. Le résultat des expérimentations menées par différents auteurs (voir Encart 8-2), conduit à une estimation de la densité de valeurs à représenter comprise entre 0,3 et 0,5 % du total des liens. Notre propre expérience sur différentes matrices des flux de l'ex-Tchécoslovaquie (1990) nous a conduites à une densité de l'ordre de 0,3%²⁰⁴.

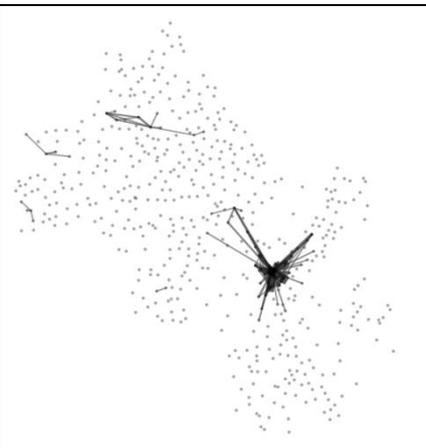
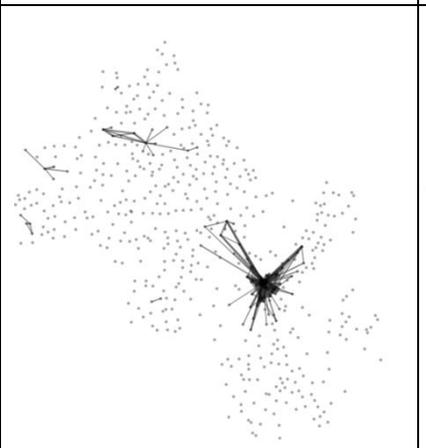
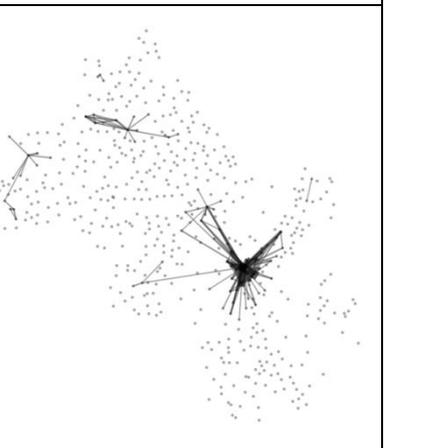
Encart 8-2. Exemples de détermination empirique du critère de densité de la matrice

Dans les cartes des migrations interdépartementales de 1954 que J. Bertin (1973) réalise à l'aide du procédé du triangle (que nous avons mentionnées page 194), l'auteur indique avoir appliqué deux restrictions : seules les valeurs de tendance migratoire supérieures à 2% et plus de la population d'origine sont représentées ; pour les autres, elles seront représentées dès lors qu'elles concernent plus de 10 000 migrants. W. Tobler estime à 0,5%, la part de liens à représenter sur une carte de flux. L'auteur illustre l'efficacité de ce seuil sur une matrice de migrations internes réalisées entre 1965 et 1970, entre les 3 141 comtés américains ($3141 * 3141$). Cette matrice est composée de 9 862 740 flux à représenter, si on élimine la diagonale. Sachant que la matrice possède une série de valeurs non disponibles, imputables à différents effets liés à la mesure des flux, seules 540 022 valeurs de flux sont renseignées. La sélection de 0,5% des flux revient alors à ne représenter que 27 003 figurés au lieu des 9 862 740. Cette valeur de densité de 0,5% correspond également à la suggestion de C. Grasland (2010), obtenue sur deux matrices de flux commerciaux : l'une à l'échelle mondiale ($181*181$ pays) en 2007, l'autre à l'échelle des 29 pays de l'OCDE, en 2005. Le raisonnement mené par D. Dorling (1991) conduit à un seuil similaire de 0,4%, obtenu par l'analyse d'une matrice de navettes interdistricts britanniques réalisées en 1981. A noter, d'une part, que ce seuil a été estimé sur une présélection des flux qui sont supérieurs à 1 000 individus ($F_{ij} > 1\ 000$), exprimés en termes de probabilités de fréquences conditionnelles qui correspondent à la pondération des valeurs de flux par les populations des lieux d'origine et de destination. D'autre part, que ce même seuil a conduit l'auteur à proposer son indice de probabilité moyenne ou globale de se déplacer.

Les résultats de l'application d'un objectif de densité des liens valués, qui correspond à la part de liens disponibles utiles, pour cartographier les ($M_{ij}+$) de l'Isère sont présentés sur la Figure 8-9 : les motifs de ces flux obtenus aux seuils de 3, de 4 et de 5% du total des liens disponibles y sont comparés.

²⁰⁴ Elle est mise en œuvre dans les cartes de flux de l'ex-Tchécoslovaquie, qui sont présentées dans la première partie, en particulier dans la section 3.3.

Figure 8-9. Cartographie des flux de l'Isère, selon différents objectifs de densité de la matrice*Matrice (F_{ij}^+)*

3%	4%	5%
 <p>40,75% des flux 3 % des liens ($\alpha=399$)</p>	 <p>44,95% des flux 4 % des liens ($\alpha=309$)</p>	 <p>48,54% des flux 5 % des liens ($\alpha=273$)</p>

Ainsi, au seuil de 3% des figurés, l'image décrit 40,7% de l'information totale, avec un critère ($\alpha=399$). On observe que ces motifs diffèrent peu lorsque la part de liens est portée à 4 ou à 5%. Ce résultat suggère, comme le précédent, le fait que le critère s'inscrit dans un continuum de valeurs. A noter que le raisonnement mené selon la part de liens conduit à une élimination des valeurs plus importante, que lorsqu'il est mené sur la part de flux.

La Figure 8-9 ne présente que les flux qui sont polarisés par les zones d'emploi de l'Isère. Elle démontre aussi le fait que la détermination d'une valeur exacte du seuil est une chimère. Ce qui importe, encore une fois, c'est bien l'objectif de la carte et l'interprétation qui en faite, en termes de flux ou de mouvement, comme l'a illustré la partie haute de la Figure 8-8 (page 338) qui a comparé 80% de l'information de flux et 80% de l'information de mouvement.

Plutôt que de refaire l'exercice de sélection des flux sur une autre matrice que celle de l'Isère, nous proposons de présenter dans la dernière section de ce chapitre, un prototype d'interface interactive en ligne qui permettrait le choix aisé d'un critère de sélection des flux à représenter.

8.4 Vers le développement d'outils informatiques : suggestion de programmes pour Rstudio©

Nous présentons dans cette section, les éléments permettant de réaliser la majorité des cartes de flux présentées dans cette section, à l'aide de RStudio.

L'objectif de cette section est, d'une part, de permettre la reproductibilité²⁰⁵ (de la majorité) des traitements réalisés dans le cadre de cette thèse et d'autre part, de proposer directement aux développeurs d'applications les grandes lignes d'une interface interactive qui permettrait de sélectionner de manière robuste, en utilisant différentes méthodes, les valeurs de flux à cartographier.

²⁰⁵ Ceci en référence au nouveau cadre dit de la « recherche reproductible » concernant la libre diffusion des résultats, mais aussi des données et des codes informatiques ayant conduit à les obtenir. Un article récent paru dans Le Monde présente cet état d'esprit. Il est disponible à l'adresse suivante : http://www.lemonde.fr/sciences/article/2013/07/15/pour-une-recherche-reproductible-publiez-vos-codes-et-donnees_3447825_1650684.html [lien vérifié le 10/04/2015].

Elle ne concerne que la partie de la production cartographique liée au traitement des données numériques, les aspects purement graphiques liés à l'esthétique de la figure devant de notre point de vue demeurer dans un logiciel de graphisme. Nous présentons dans un premier temps l'interface proprement puis proposons en guise de contribution une série de programmes développés pour la plateforme RStudio© qui pourront servir de point de départ.

8.4.1 Vers le développement d'une interface interactive

L'interface permettrait, en effet, d'enchaîner la chaîne de cartographie d'une sélection de valeurs de flux, elle est présentée sur la Figure 8-10.

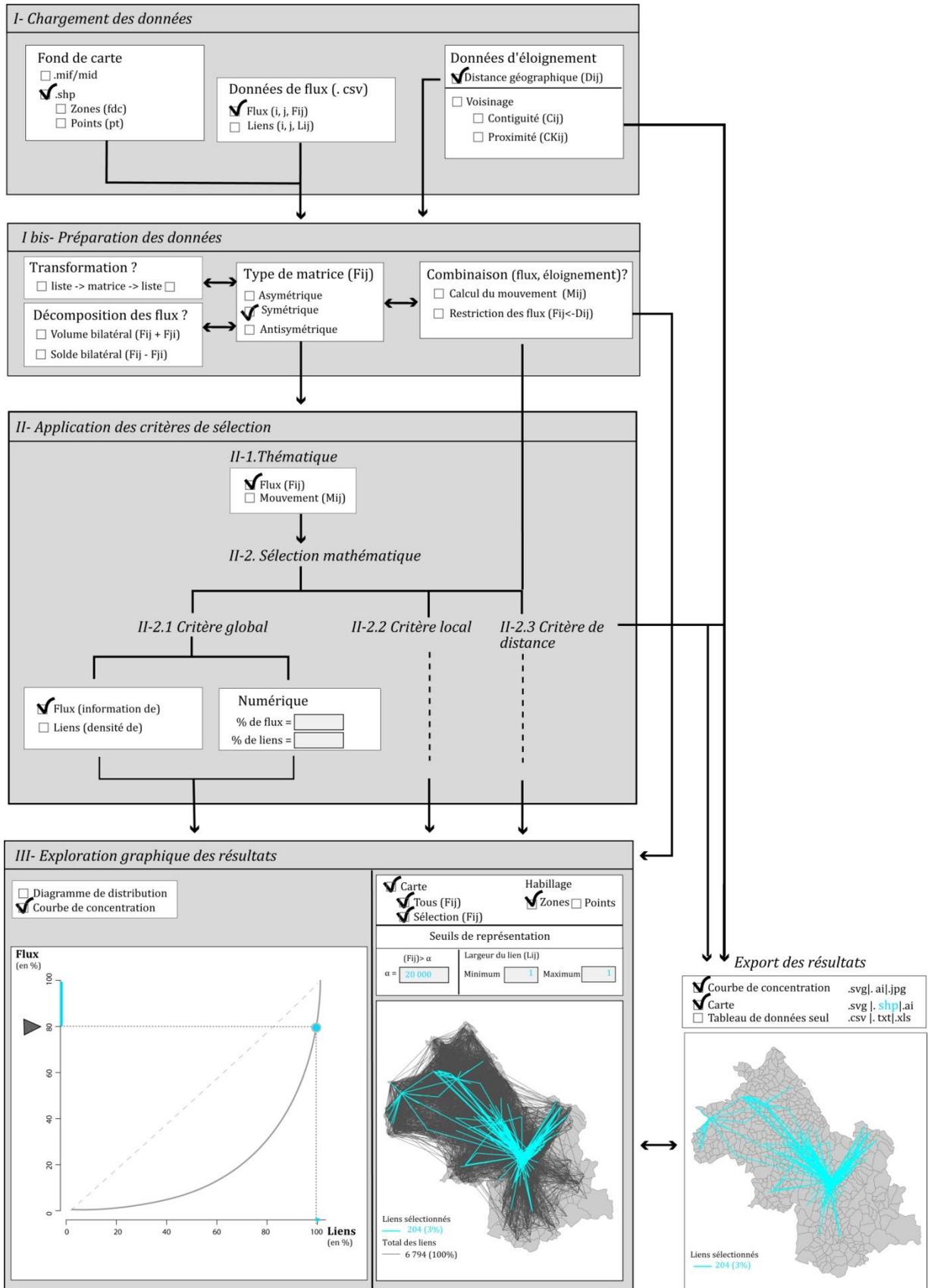
Cette proposition d'application est générale, elle est composée de quatre à cinq étapes principales qui pourraient être enchaînées de la manière suivante : Chargement des données → (facultatif : préparation des données) → Application des critères de sélection (thématique, graphiques ou numériques) → Tracé de la courbe de concentration (et/ou du diagramme de distribution) et choix du double critère de sélection ↔ Affichage des résultats de la sélection sur un fond de carte → Exportation des résultats (sous la forme d'un fichier numérique tableau ou cartographique).

La première étape consisterait dans le chargement des données cartographiques et numériques sur les flux, dans des formats permettant de gérer l'interopérabilité de l'interface avec les outils de traitement statistique (R, par exemple) et de géomatique. Le format proposé pour le fond de carte est classique : il s'agit d'une couche vectorielle formé du maillage surfacique de la zone d'étude, naturellement associée à une table attributaire : il s'agit du *shapefile* de ESRI©, et du format d'échange mif/mid de Mapinfo©. D'autres formats peuvent être proposés, parmi lesquels le JSON

Pour les données numériques, le format idéal est le (.csv). Les deux tableaux suivants sont *a minima* requis : des données de flux (origine, destination, F_{ij}), présentés de préférence sous la forme de liste (voir chapitre 1) et la liste des identifiants des N lieux, avec leurs coordonnées géographiques (ID , i , x_i , y_i). Un troisième fichier, portant sur l'éloignement des lieux (voir chapitre 6) sera requis, dans l'éventualité où l'on souhaiterait réaliser une cartographie du mouvement. Si tel est le cas, la possibilité serait alors offerte de générer une matrice de mouvement (M_{ij}), issue de la combinaison entre (F_{ij}) et (D_{ij}) lors de l'étape facultative de préparation des données. A ce stade, il serait également proposé, lorsque la matrice initiale (F_{ij}) est asymétrique, de la décomposer en matrice du volume (F_{ij}^+) ou du transfert net bilatéral (F_{ij}^-), comme présenté dans le chapitre 3.

La seconde étape consisterait à appliquer les différents critères de sélection : thématiques puis graphique (global ou local). Dans le premier cas, il s'agit de décider de l'interprétation qui sera donnée aux données analysées et *in fine* à la carte (de flux ou mouvement). La possibilité d'interprétation en termes de mouvement ne sera offerte que dans le cas où une matrice d'éloignement aura été sélectionnée lors de l'étape précédente et cela, afin d'éviter des erreurs d'interprétation. Dans le second cas, le choix du critère de sélection serait soit global (comme présenté dans le présent chapitre 8, soit local, il concernera alors les propositions présentées dans le chapitre 9). Un même critère global pourrait être déterminé de trois façons différentes : en fonction de l'objectif de la sélection (de significativité ou de densité, voir section 8.3), de façon numérique (en indiquant la part de liens ou la part d'information à représenter) ou, lors de l'étape suivante, de manière graphique. A ce stade, l'utilisateur raisonne *a priori* pour déterminer le critère de sélection (global) de flux, en l'occurrence en fonction des objectifs de *significativité* ou de *densité* de la matrice. Il détermine alors soit le % de flux, soit le % de liens à cartographier, tous deux étant rapportés à l'ensemble de l'information. Cette possibilité de détermination du critère de sélection doit rester ouverte, afin de satisfaire aux différentes approches et logiques du raisonnement sous-jacent.

Figure 8-10. Architecture générale d'une application en ligne de sélection interactive de l'information de flux à cartographier



Françoise Bahoken, 2015

La troisième étape viserait à explorer les résultats issus de l'application du critère de sélection, sur différents types de graphiques, en particulier sur la courbe de concentration des valeurs de flux ou sur la couche cartographique de l'ensemble des liens. Cette étape de co-visualisation graphique et cartographique est cruciale : elle permet d'explorer le résultat issu de l'application d'un critère de sélection des valeurs, par rapport à l'ensemble de l'information de flux (ou des liens) disponible(s) : aussi d'ajuster éventuellement ce seuil ; et de représenter les flux, en paramétrant les seuils graphiques de perception des flux sur une carte (les valeurs minimales et maximales de la *largeur* du figuré).

Sur la Figure 8-10, nous avons mentionné en bleu, les aspects liés à l'interactivité de l'interface : il concerne le choix graphique (visuel) du critère de sélection et d'autre part, l'affichage simultané du résultat cartographique, et de la valeur (α) du critère. C'est à ce stade que la possibilité de sélection graphique est offerte. L'observateur dispose alors de la possibilité de déplacer un curseur placé sur la courbe de concentration, ou sur l'axe des ordonnées (pour agir sur la part en % de flux) ou sur celui des abscisses (pour la part en % de liens) ; l'action sur l'un de ces trois curseurs actionnant automatiquement les autres. Dans l'exemple présenté sur la Figure 8-10, l'utilisateur agit sur la part en % de flux. Cette approche visuelle est complémentaire de l'approche numérique de la sélection proposée à l'étape précédente.

L'affichage du résultat étant simultané (avec ou sans fond de carte), l'utilisateur peut ensuite faire le choix d'exporter les différents résultats obtenus : le graphique de la courbe de concentration, la sous-matrice des valeurs de flux sélectionnées (.csv), la couche vectorielle (.shp) correspondant aux valeurs de flux sélectionnées. A noter qu'il serait intéressant d'exporter une modification de la table de départ, en ajoutant une ou plusieurs colonnes de résultats de la sélection.

L'illustration de la manière dont cette interface pourrait fonctionner a été mise en œuvre sur la matrice de flux (Mij+) de l'Isère (voir Figure 8-7). Celle-ci pourrait être développée sur la plateforme RStudio avec RShiny et pourrait, à terme faire également l'objet d'un package dédié à l'analyse des flux. Une application a d'ailleurs été initiée dans cette perspective.

8.4.2 Suggestion de programmes RStudio©, de sélection de flux à cartographier selon un critère global

Nous avons vu dans le chapitre 2 que l'analyse (carto)graphique d'une matrice de flux pouvait être menée selon deux logiques de raisonnement : l'une matricielle et l'autre réticulaire. Ces deux logiques correspondent à deux modèles graphiques de représentation de la même information, sous les formes respectives de tableau croisé et de graphe valué spatialisé. Elles correspondent à deux formalisations conceptuelles, la première mobilisant l'objet matrice – selon l'approche que nous avons privilégiée dans cette thèse – et la seconde l'objet graphe et la théorie des graphes sous-jacente. De ce point de vue, l'étape de préparation et de traitement des données de flux peut être réalisée dans l'un ou l'autre environnement conceptuel et conduire à quelques nuances près strictement au même résultat.

Nous présentons en annexes une série de cinq programmes réalisés sous RStudio :

- *l'Annexe R-1 : Procédure de transformation d'une matrice de flux fermée non carrée en une matrice de flux carrée* où le nombre de lignes est égal au nombre de colonnes : cette opération est rendue nécessaire pour pouvoir ensuite décomposer la matrice. L'exemple d'illustration est la matrice d'origine de l'Isère ;
- *l'Annexe R-2 : Procédure de cartographie d'une matrice de flux* présente un enchaînement de commandes permettant de cartographier directement des flux, à partir de données qui sont d'une part, numériques (au format matriciel ou liste) que l'on décompose ou non ; et d'autre part, cartographiques (fond de carte zonal au format *shapefile*). Cette annexe est transversale et l'ensemble des chapitres ;

- *l'Annexe R-3 : Procédure de sélection des flux à cartographier, selon un critère spatial (de voisinage binaire, ordinal ou de distance parcourue)*, est composée de cinq sections qui conduisent à sélectionner les flux selon des critères liés au voisinage spatial binaire (C_{ij}) ; ordinal (CK_{ij}), des k plus proches voisins ; et selon la distance parcourue (D_{ij}), incluant une réduction de (F_{ij}) en fonction de (D_{ij}). Ces sections correspondant au chapitre 7. Le cas d'illustration est la matrice de flux de navettes de l'Isère ;
- *l'Annexe R-4 : Procédure de sélection des flux à cartographier selon un critère global*, considère d'abord un critère unique ($F_{ij} > \alpha$) interne ou externe ($F_{ij} \leftarrow (D_{ij}) > \alpha$) de réduction de la matrice, puis un double critère de significativité de la part de liens en fonction de la part de liens (F_{ij} , L_{ij}), en s'appuyant sur un ensemble de programmes réalisés avec C. Grasland (2010, 2014) ; cette annexe correspondant aux solutions présentées essentiellement dans le chapitre 8. Le cas d'illustration est la matrice de flux domicile-travail de l'Isère ;
- *l'Annexe R-5 : Procédure de sélection des flux à cartographier, selon un critère local (« flux dominants »)* permet d'appliquer les critères de sélection proposés par la méthode Nuysten-Dacey (1961) présentés dans le chapitre 9, ainsi que quelques variantes ;
- *l'Annexe R-6 : Procédure de cartographie directe d'une variable de l'environnement global de Rstudio (c)* pour optimiser la production de carte de flux à partir d'une série de matrices préalablement construites.

Cette série de programmes est réalisée selon l'approche matricielle que nous avons privilégiée tout au long de cette thèse. Elle sera en ligne, déposée sur le site du Carnet de recherche du groupe de travail ElementR. Cette page dédiée est accessible à l'adresse suivante : <http://elementr.hypotheses.org/314> [lien inséré post-soutenance, le 03/02/2016]. Toutefois, il convient de noter que certaines fonctions des chapitres 7 et 8 mettent en œuvre une logique réticulaire, dont une première ébauche avait été réalisée en collaboration avec P. Chapron : elle est également disponible en ligne, sur le site de ElementR²⁰⁶. Si l'intérêt de cette seconde version porte sur son optimisation au niveau du codage, son principal défaut est de ne pas autoriser l'export des matrices intermédiaires et de n'être pas suffisamment avancée au niveau du rendu graphique. Mais tel n'était pas l'objet initial de cette suite de programmes. C'est pourquoi l'approche matricielle peut apparaître plus aboutie.

²⁰⁶ Le *Markdown* de l'approche réticulaire de sélection globale des flux peut être téléchargé à l'adresse suivante : <http://elementr.hypotheses.org/204> [Lien vérifié le 10 avril 2015].

Conclusion du chapitre

Les solutions liées à l'application d'un *critère global unique* de sélection des flux à cartographier s'inscrivent dans le cadre d'un raisonnement qui consiste à estimer la part *d'information significative* de flux à représenter sur la carte. La particularité de l'approche est de ne mobiliser qu'un et un seul critère, qui agit sur l'ensemble de la matrice. Elle conduit à la réalisation d'une carte unique et non d'une collection de cartes. L'approche diffère en ce sens de celle que nous avons examinée dans le chapitre 7, qui consistait à déterminer plusieurs critères à partir desquels la matrice de flux serait segmentée.

L'approche proposée a, en effet, consisté à dérouler le fil d'une idée suggérée par W. Tobler (1982) d'examiner la courbe de distribution des valeurs de flux. Cela nous a conduits à mobiliser dans un premier temps, une courbe de concentration, à la place de celle des fréquences simples. Nous montrons qu'elle agit en fait sur un *double critère unique de sélection* qui correspond à la part d'information de flux *significative* et à la part de liens représentés sur la carte. La méthode fondée sur l'examen de la courbe de concentration des flux permet d'estimer visuellement la valeur du critère qui agit sur les plans analytiques et graphiques. Ce critère s'inscrit soit, dans un intervalle de valeurs de flux concentrées, ce qui correspond à la sélection des plus gros flux ; soit, et à l'inverse, dans un intervalle de dispersion des liens, qui correspond à la réalisation d'une carte qui présente peu de figurés mais qui sont représentatifs de l'information disponible dans la matrice. L'enjeu étant alors celui de l'arbitrage entre une part importante de l'information représentée et une part faible de liens.

Si l'approche proposée a été illustrée sur une matrice du volume bilatéral de flux (F_{ij}^+), son caractère généralisable la rend applicable à l'ensemble des familles de matrices, en utilisant uniquement les valeurs positives dans le cas de la matrice (F_{ij}^-).

L'une des particularités de l'application d'un *critère unique de sélection*, que nous avons mentionnée en introduction, tient à l'élimination de certains lieux de l'analyse. Il s'agit de ceux dont les flux (entrant ou sortant) ne répondent pas au critère de sélection. Cependant, le fait de considérer certains lieux comme inintéressants peut être considéré comme gênant pour l'analyse. Si tel est le cas, il est possible d'appliquer un critère de sélection qui conserve l'ensemble des lieux de la zone d'étude, en agissant au niveau local de la matrice, sur les lignes et/ou les colonnes de la matrice ; les solutions correspondantes font l'objet du chapitre 9.

9^{ième} CHAPITRE. Propositions de solutions liées au choix d'un critère de sélection local : la cartographie de flux dominants

« L'objectif de ce document est de décrire une procédure d'ordonnement et de regroupement de villes en fonction de l'intensité et de la direction de flux de marchandises, de personnes, et de leurs relations. Les théories actuelles des régions nodales et des places centrales fournissent les bases à la mise en évidence d'une organisation hiérarchique régionale des réseaux de villes ».

John D. Nyusten et Michael F. Dacey (1961).

Introduction du chapitre

Ce neuvième chapitre a pour objectif de présenter les solutions liées au choix d'un *critère de sélection local* à appliquer aux cellules de la matrice de flux, au préalable de leur cartographie. Cette procédure est rendue nécessaire lorsque la matrice est dense et que l'objectif n'est pas de représenter, d'une part, toute *l'information disponible* qu'elle contient (des solutions sont proposées dans le chapitre 5) et d'autre part, une sélection issue d'une procédure de filtrage global (voir chapitre 8) ou mobilisant un critère de distance parcourue ou de voisinage (voir chapitre 7).

A la différence des précédentes propositions, le choix d'un *critère de sélection local* focalise l'analyse sur les lieux et non sur les liens : il mobilise les valeurs de flux (Fij) pour caractériser les N lieux de la zone d'étude en termes d'interactions préférentielles mettant en évidence une organisation hiérarchique. Le résultat de cette procédure conduit, en effet, à la réalisation d'une carte de flux remarquable par rapport à celles que nous avons analysé jusqu'ici, car ses motifs révèlent des relations hiérarchiques entre des couples de lieux, définissant des *régions nodales* autour des lieux les plus attractifs. Ces régions sont organisées par plusieurs pôles nodaux dits dominants, elles définissent le périmètre d'influence de ces derniers et décrivent l'arborescence des relations qui caractérisent les localités concernées.

L'approche proposée revêt deux particularités notables. Elle s'inscrit dans le cadre conceptuel de la théorie des graphes, mobilisant plutôt une matrice de type (Lij) que (Fij) et elle est focalisée sur le poids des lieux (Oi ou Dj dans l'ensemble des interactions), pour réaliser une cartographie conjointe liens-lieux.

La théorie des graphes constitue effectivement un cadre d'analyse intéressant des matrices de type (Lij), qui correspondent formellement à des *matrices d'adjacence* (voir chapitre 1). Sa mise en œuvre dans un objectif de cartographie présentant de nombreux intérêts qui auraient pu justifier l'inscription de l'ensemble de notre analyse dans son cadre conceptuel. Elle procure une batterie d'indicateurs et de mesures variées permettant de caractériser et décrire les graphes - que l'on aurait pu considérer valués, éventuellement spatialisés -, ce d'autant plus que nous avons montré la pertinence du graphe pour représenter des flux (chapitre 6), lorsque la composante spatiale n'a pas d'intérêt pour l'analyse.

En recourant à la théorie des graphes pour examiner les possibilités de sélection des valeurs de flux à représenter, nous changeons ponctuellement de cadre théorique de référence. Ce changement de

posture ne conduit pas, pour autant, à une réorientation soudaine de notre travail. Nous allons en fait nous servir de la grille d'analyse des réseaux sociaux que cette théorie procure, pour enrichir l'analyse cartographique de flux. Pour cela, nous allons supposer que la matrice (Fij) correspond à un réseau non planaire de (i) acteurs localisés et inter reliés. Chacun d'entre eux est caractérisé par un degré d'incidence ou un degré pondéré « [...] correspondant au nombre de liens ayant le sommet comme extrémité (si le lien est une boucle, il est compté deux fois). La mesure du degré des différents sommets/acteurs est un des moyens utilisés pour repérer les sommets/acteurs les plus importants dans un graphe [...] Lorsque le graphe est orienté, on distingue le degré entrant (in-degree) qui correspond au nombre de liens se dirigeant vers le sommet et le degré sortant (out-degree) qui correspond au nombre de liens partant du sommet »²⁰⁷ (Drevelle, 2013).

Nous allons ainsi analyser l'ordre (k) des degrés sortants de ces acteurs (des lieux) autrement dit, le rang des valeurs de flux (Fij) pris dans la distribution d'ensemble des valeurs de chacun des lieux (i). A la différence du chapitre 8, nous faisons le choix de conserver l'ensemble des lieux de la zone d'étude soit au moins un flux pour chaque lieu d'origine (ou de destination). Nous posons dès lors pour tout (i), (k≥1).

L'introduction de (k≥1) nous place devant un problème qui consiste à réduire la densité de (Fij), correspondant à [N*(N-1)], en une matrice à (k*N) valeurs renseignées, avec :

(N) : le nombre de lieux ;

(k : 1, ..., n) : le nombre de liens (sortants ou entrants) par lieu.

Nous examinerons ainsi, dans un premier temps, le cas de (k=1). La matrice de départ étant asymétrique, elle est formée de $\frac{N(N-1)}{2}$ valeurs, soit (N-1) liens à représenter. Etant donné l'étendue des possibilités de sélection des valeurs de flux émises selon leur rang, nous allons considérer la valeur la plus importante $Max_i (Fij)$, correspondant formellement au *flux majeur*. Ce choix n'est pas anodin : il permet de nous inscrire dans le cadre d'analyse de J.D. Nyusten et M. F. Dacey (1961), déjà mentionné (voir section 2.2.2.1, page 70), qui considère une acception du flux le plus important pour réduire la matrice de flux en une matrice d'adjacence localisée. « [...] La relation de type 'flux le plus important' [*largest flow*] peut avoir plusieurs définitions, comme celle du plus large flux sortant [Fij], entrant [Fji] ou total [que l'on supposera correspondre à [Fij⁺]] » (Nyusten et Dacey, 1961: 33).

De fait, l'intérêt du recours à l'approche de J.D. Nyusten et M. F. Dacey (1961) est multiple.

Le premier intérêt porte sur la définition de régions nodales issues du système de flux, à la suite de l'application de deux conditions de sélection qui agissent respectivement sur les liens et sur les lieux. Cela tient dans l'adoption d'emblée d'une grille d'analyse conjointe des liens et des lieux, dans un processus (carto)graphique. Cette posture implique de prendre en considération l'ensemble des parties de la matrice : de ses valeurs ainsi que des aspects de sémiologie cartographique qui en dépendent à savoir la représentation des liens ordinaires (Lij) ou valués : (Fij) ou (Mij), mais également des lieux (i) et cela, pour la première fois depuis le début de cette thèse. En effet, aucune des méthodes présentées dans cette seconde partie n'a focalisé l'attention simultanément sur les liens et les lieux, sur le cœur de la matrice et ses marges à la fois.

Ainsi, si le cadre d'analyse de J.D. Nyusten et M. F. Dacey (1961) permet d'enrichir l'analyse cartographique de flux, il impose de considérer les problèmes spécifiques découlant d'une analyse réalisée du point de vue des lieux, plus spécifiquement l'*effet de taille des lieux* (dont nous verrons qu'il indique le caractère dominant ou non d'un lieu, ou bien la structure arborescente de l'ensemble)

²⁰⁷ Ce billet a été rédigé en 2013 par M. Drevelle, il est disponible à l'adresse suivante : <http://groupefmr.hypotheses.org/1425> [lien vérifié le 6/08/2015].

et de convergence²⁰⁸ des flux qui lui est traditionnellement associé, ces derniers en lien avec l'effet spaghetti.

Le second intérêt de ce cadre d'analyse tient à la variété de résultats envisageables, que nous nous proposons d'examiner sur le plan cartographique. En effet, nous conservons notre objectif initial de réduction de la densité de la matrice pour produire une carte de flux lisible et compréhensible en mobilisant des critères locaux.

Nous présentons, dans une première section, l'approche classique dite Nyusten-Dacey, telle que ses auteurs la mettent en œuvre par l'exploitation d'une matrice de flux (Fij) téléphoniques entre une sélection de villes américaines. La méthode étant ancienne et, *a priori*, bien connue, on ne s'attend probablement pas à sa présentation exhaustive. Nous la valorisons ici pour trois raisons principales. La première raison est liée au résultat obtenu suite à la sélection de valeurs de flux opérée localement : elle conduit *in fine* à une perception de la structure hiérarchique des lieux. La seconde raison concerne la variété de résultats (carto)graphiques envisageables qui sont rarement mobilisés simultanément. La troisième raison tient au fait que la version initiale de l'approche ne soit pas couramment mise en œuvre et demeure, de fait, mal connue. Elle conduit pourtant à une double représentation des liens (Lij) – et non de (Fij) – et de l'attractivité des lieux (Dj), éludant ainsi la valeur des uns pour valoriser les autres. C'est, par exemple, le cas des cartes de flux dominants réalisés par N. Cattan (2004), déjà mentionnées dans le chapitre 7, qui ne mobilisent qu'une partie de la méthode, de même celles de C. Ducruet²⁰⁹.

Dans la seconde section, nous proposons une (re)construction cartographique de la méthode Nyusten-Dacey, en examinant les cinq étapes qui la constituent. La première correspond à la cartographie des flux majeurs $\text{Max}_i(\text{Fij})$. Elle est réalisée selon l'approche usuelle, en agissant sur la *largeur* d'un figuré linéaire (un trait droit, par exemple), en application de la *Taille*. La seconde étape introduit la notion de flux *dominant* dans l'analyse, conduisant à une modification du raisonnement caractérisée par le passage d'une *logique de flux* à une *logique liens-lieux*. L'analyse est désormais focalisée sur la valeur des lieux (Dj) et non plus sur celle des liens (Fij). Sur le plan cartographique, cela a pour conséquences que les traits valués ne le sont plus, ils sont remplacés par des traits d'épaisseur constante $(\text{Lij}) \leftarrow (\text{Fij})$. L'analyse des lieux mobilise, quant à elle, une implantation ponctuelle et la *Forme* (un cercle, par exemple) pour les symboliser. La représentation de la variable quantitative (Dj), associée aux points, permet de caractériser leur attractivité au regard du critère de dominance impliquant *le flux majeur* $\text{Max}_i(\text{Fij})$ et $(\text{Dj}) > (\text{Di})$; elle implique aussi la *Taille*²¹⁰ pour distinguer les lieux dominants de ceux qui n'en dominent aucun. La troisième étape consiste à combiner les deux précédentes, en gérant les positions dominant et dominé à la fois, ou intermédiaires des lieux. Pour cela, elle précise leur caractérisation en utilisant, de nouveau et principalement²¹¹, la *Forme* (des cercles et des carrés, par exemple). La quatrième étape est celle de la régionalisation : elle consiste, d'une part, à définir l'aire d'influence des lieux dominants, à distinguer les différents bassins des lieux dominants en recourant à la *Couleur*. D'autre part, à identifier des sous-bassins, en mobilisant la *Valeur* pour les caractériser au regard de la position des lieux dominants ou intermédiaires. Le cas d'application proposé est la matrice de l'Isère (Fij) pour laquelle nous examinons différentes possibilités de représentation de l'attractivité et la hiérarchie des lieux puisque la méthode est avant tout locale.

²⁰⁸ Nous rappelons que cet effet résulte du fait que les flux les plus importants se dirigent vers les zones de plus forte taille (en valeur), mais qui souvent les plus petites en surface, en raison de la manifestation de différents phénomènes d'accumulation.

²⁰⁹ <http://www.lboro.ac.uk/gawc/rb/rb364.html> [lien vérifié le 02/08/2015].

²¹⁰ La *Taille* peut également servir ici à représenter l'attractivité des lieux (Dj).

²¹¹ En fait, il est possible de mobiliser la *Forme*, l'*Oriente* et la *Couleur* mais seule la première variable visuelle sera considérée dans cette troisième étape, la couleur étant réservée pour la quatrième étape.

La troisième section a pour objectif de valider les solutions proposées sur une matrice de flux commerciaux internationaux (OMC, 2007). Elle vise à examiner les conditions de généralisation de la méthode en présentant, dans un premier temps, le cas usuel d'application de la méthode Nyusten-Dacey sur une matrice (Fij). Son extension à deux matrices complémentaires (Fji et Fij⁺), correspondant aux possibilités théoriques et pratiques, est envisagée dans un second temps. Le cas de (Fij⁻) est d'emblée éliminé, car il pose un certain nombre de problèmes thématiques et pratiques qu'il n'est pas possible de résoudre à ce stade. L'examen conjoint de (Fij) et (Fji), éventuellement de (Fij⁺), permet d'enrichir la méthode en considérant la réversibilité des conditions de sélection des valeurs de flux, autrement dit le fait que les poids des lieux d'origine ne sont pas forcément identiques aux poids des lieux de destination.

9.1 La méthode Nyusten-Dacey (1961) d'analyse des *flux dominants*

La méthode d'analyse dite des flux *dominants* a été développée au début des années 1960 par J. D. Nyusten et M. F. Dacey. Elle s'inscrit suite à la démonstration selon laquelle « [...] *l'orientation et l'intensité des flux liés aux processus sociaux sont des indicateurs de l'organisation spatiale de la structure régionale de la société urbaine* » (Nyusten, Dacey, 1961 : 29). En effet, les auteurs sont, à l'époque, à la recherche de solutions à l'identification d'un problème de régionalisation des échanges autour de lieux attractifs, de création de régions nodales (Nyusten et Dacey, 1961 ; Dumolard, 1975) ainsi structurées par un système de flux nodaux.

La méthode a ensuite été reprise et appliquée en France (Dumolard, 1975 : 107), grâce à la traduction française de l'ouvrage de P. Haggett (1973) qui favorisa son appropriation immédiate. P. Dumolard suggère d'ailleurs d'utiliser le terme « *d'espaces polarisés* » qui sont ainsi structurés autour d'un pôle, par un système de flux nodaux, que l'usage francophone qualifie de *flux dominants*.

9.1.1 Présentation des données (de la matrice disponible pour l'analyse)

J. D. Nyusten et M. F. Dacey présentent, en 1961 lors d'un colloque de Science régionale, l'analyse d'une matrice de flux téléphoniques, à longue distance, réalisés entre 40 villes américaines de l'état de Washington. Dans les actes associés à cette communication (Nyusten et Dacey, 1961 : 38-42), les auteurs illustrent en partie leur exposé par une matrice fictive présentée sur la Figure 9-1.

Cette matrice est de type (Fij) à diagonale vide et de rang 12. Elle décrit une quantité de communications téléphoniques réalisées pendant une semaine de juin 1958. Cette matrice est mobilisée dans l'objectif de décrire « [...] *la procédure d'ordonnement et de regroupement de lieux en fonction de l'intensité et de l'orientation de flux de marchandises, de population ou de communication entre eux. [...] Le problème consiste à développer une méthode capable de quantifier le degré d'association de paires de lieux de sorte qu'il permette d'identifier les associations les plus fortes du réseau. Ces actions peuvent être exprimées en termes d'interactions directes entre deux villes, ou indirectes par l'intermédiaire d'une ou de plusieurs villes. L'intensité de l'association directe ou indirecte est mesurée par un indice dérivé de la théorie des graphes* » (Nyusten et Dacey, 1961 : 29).

Les auteurs préparent ainsi les données de cette matrice (avant sa représentation graphique) en suivant une procédure formée de trois étapes que nous avons reconstituées et présentons ci-après.

Figure 9-1. La matrice de flux téléphoniques de (Nyusten et Dacey, 1961)

		TO CITY											
		a	b	c	d	e	f	g	h	i	j	k	l
FROM CITY	a	0	75	15	20	28	2	3	2	1	20	1	0
	b*	69	0	45	50	58	12	20	3	6	35	4	2
	c	5	51	0	12	40	0	6	1	3	15	0	1
	d	19	67	14	0	30	7	6	2	11	18	5	1
	e*	7	40	48	26	0	7	10	2	37	39	12	6
	f	1	6	1	1	10	0	27	1	3	4	2	0
	g*	2	16	3	3	13	31	0	3	18	8	3	1
	h	0	4	0	1	3	3	6	0	12	58	4	0
	i	2	28	3	6	43	4	16	12	0	58	13	1
	j*	7	40	10	8	40	5	17	34	98	0	35	12
	k	1	8	2	1	18	0	6	5	12	30	0	15
	l	0	2	0	0	7	0	1	0	1	6	12	0
	Column Total		113	337	141	128	290	71	118	65	202	311	91

Largest flow circled.
Largest flow determined by the number of out-going messages.

*Largest flow from these cities is to a "smaller" city where "size" is determined by the column totals.

Source : J. D. Nyusten et M. F. Dacey, 1961 : 35

La première étape du traitement des données correspond à l'identification du « flux le plus important » (*largest flow*), traduit en français par l'expression *flux majeur*, défini comme suit par les auteurs. « La relation de type 'flux le plus important' peut avoir plusieurs définitions, comme celle du plus large flux sortant » (Nyusten et Dacey, 1961: 33). Celui-ci correspond à la valeur maximale du flux $[F_{ij} \leftarrow \text{Max}_i(F_{ij})]$ émise par chaque ville, c'est-à-dire au flux de rang 1 et qui est dirigé vers une ville de plus grande taille.

Ce flux maximal ou *flux majeur* est défini au regard du seul lieu d'origine (i), de la manière suivante

Définition du flux majeur :

Soit (F_{ij}) , la valeur du flux entre un couple de lieux d'origine et de destination (i,j) .

On pose (m) le lieu de destination du flux majeur (F_{im})

avec :

$(m) : 1, \dots, N$

N : nombre total de lieux de destination.

(F_{im}) est un flux majeur si et seulement si $(F_{im}) = \text{Max}_i(F_{ij})$ [9.1]

La seconde étape du traitement porte sur la réduction de la matrice des *flux majeurs* (F_{im}) en une matrice des *flux dominants* (F_{id}) . Elle considère le *flux majeur* émis vers des zones de taille supérieure (précisée sur la légende des auteurs, située au bas de la Figure 9-1). (F_{id}) est issue de (F_{im}) réduite aux seuls flux dirigés vers un lieu plus attractif, elle est mise en œuvre de la façon suivante.

Définition du flux dominant :

Soit (F_{im}) , le *flux majeur* émis de (i) vers (m) , avec $(F_{im}) = \text{Max}_i(F_{ij})$

On pose (d) le lieu de destination du *flux dominant* (F_{id}) avec :

$(d) : 1, \dots, N$

avec :

$N = 1$ (N : nombre de lieux de la zone d'étude)

(F_{id}) est un *flux dominant* si et seulement si

$(F_{id}) = (F_{im})$ et $(D_d) > (D_i)$ [9.2]

avec :

$Dd : \sum_{j=1}^n (F_{ij})$ somme marginale en colonne de (d)

$Di : \sum_{k=1}^n (F_{ij})$ somme marginale en colonne de (i)

Les *flux dominants* correspondent ainsi aux huit valeurs entourées sur la Figure 9-1. Légendées sur la partie droite de la figure, elles sont définies du seul point de vue des lieux d'origine (en ligne). Ces huit valeurs excluent quatre villes (b, e, g, h) parmi les douze de l'exemple, et sont associées d'un astérisque : ces villes correspondent à celles émettant des *flux majeurs* qui ne sont pas des *flux dominants*, car ils sont dirigés vers des villes de taille inférieure.

La troisième étape du traitement des données porte sur la réduction la matrice (Fid) en une matrice d'adjacence (Lid), présentée en [9.3], et sur la Figure 9-2 ci-dessous :

(Lid) \leftarrow (Fid) [9.3]

Figure 9-2. La matrice d'adjacence des flux téléphoniques de (Nyusten et Dacey, 1961)

Adjacency Matrix of Graph G

	a	b	c	d	e	f	g	h	i	j	k	l	
a	1												*
b		1											
c			1										**
d				1									*
e													*
f						1							*
g													*
h													*
i									1				*
j									1				*
k										1			*
l											1		*

Blank spaces represent zero elements.

*Terminal point. **Trivial terminal point.

Source : J. D. Nyusten et M. F. Dacey, 1961 : 35.

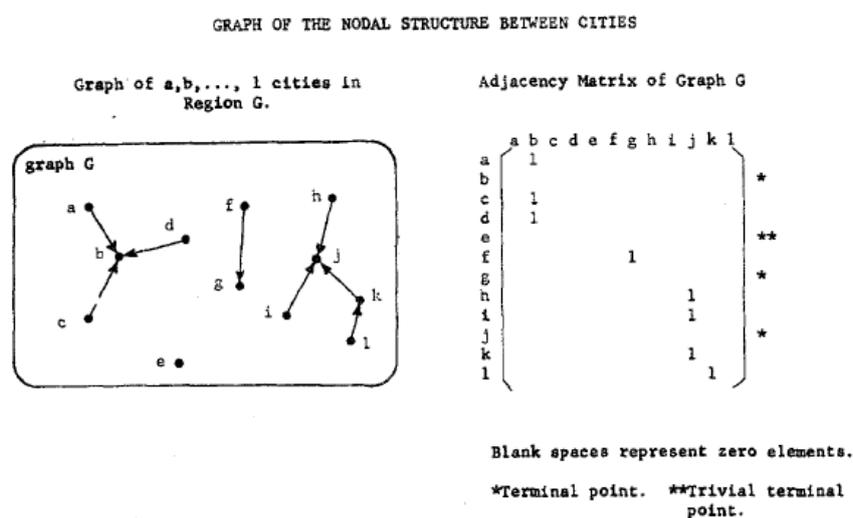
Il est important de noter que la donnée disponible pour la cartographie des *flux dominants* est une matrice qui décrit des relations (Lid) issues de l'analyse des *flux dominants* et non les *flux dominants* (Fid) en eux-mêmes, ces derniers ayant été perdus dans l'analyse. Cette matrice binaire (voir Figure 9-2) conserve toutefois l'information sur l'importance des lieux, en deux modalités principales qui sont mentionnées sur la marge en ligne par des astérisques : celle des « *points terminaux* » et celle des « *points terminaux triviaux* ». Cette mention, portée en légende de la Figure 9-2, fait référence aux caractéristiques : « *dominant* » ou « *satellite* » des lieux.

La préparation des données terminée, les auteurs se livrent ensuite à sa représentation graphique puis cartographique, nous les présentons successivement dans les deux sous-sections suivantes.

9.1.2 Représentation graphique des relations nodales

La seconde partie de l'analyse consiste en la représentation de la matrice des relations nodales (Lid) sous la forme d'un *graphe*. La Figure 9-3 décrit la sélection de liaisons téléphoniques issues des flux dominantes entre les villes de la région de Washington présentée par les auteurs.

Figure 9-3. Le graphe nodal des flux téléphoniques (Nyusten et Dacey, 1961)



Source : J. D. Nyusten et M. F. Dacey, 1961 : 35.

Si le graphe est naturellement formé de liens orientés, il n'est pas valué, de même que la matrice qui lui est associée n'est plus discrète, mais binaire (nous y reviendrons dans la section 9.1.3.). **La matrice de départ (Fij) a ainsi été réduite (aux seuls flux dominants) puis dégradée, à la seule mention de l'existence d'un lien entre deux lieux et non plus de leur valeur.** Cette action correspond d'ailleurs aux deux possibilités de cartographie de flux envisagées par J. Bertin (voir Figure 4-17, page 174).

Le graphe des *flux dominants* est formé de points et de flèches orientées (voir partie gauche de la Figure 9-3). Il décrit une structure acyclique et hiérarchique, formée de sous-composantes connexes décrivant des arbres hiérarchiques organisés autour de lieux remarquables (b, g, j, k et e) ; l'ensemble des lieux étant représenté par un point noir. D'après la partie droite de la Figure 9-3, certains des lieux remarquables du graphe sont associés d'astérisques, en fonction de leur position dans l'arbre : ils correspondent soit à « *point terminal* » (b, g, j), soit à « *point trivial* » (e), en l'occurrence un point isolé.

La troisième étape consiste en une représentation cartographique de cette structure nodale régionale, car ni le graphe ni la matrice ne permettent de figurer la géographie de ce système de relations (voir chapitre 2). Toutefois, leur mobilisation conjointe au sein de l'analyse d'une matrice unique enrichit indéniablement le propos. On voit bien ici la complémentarité de ces trois modèles graphiques dans le cadre d'analyse des flux, la proximité des résultats menés selon un raisonnement sur les interactions spatiales (chapitre 6) ou du point de vue de la théorie des graphes

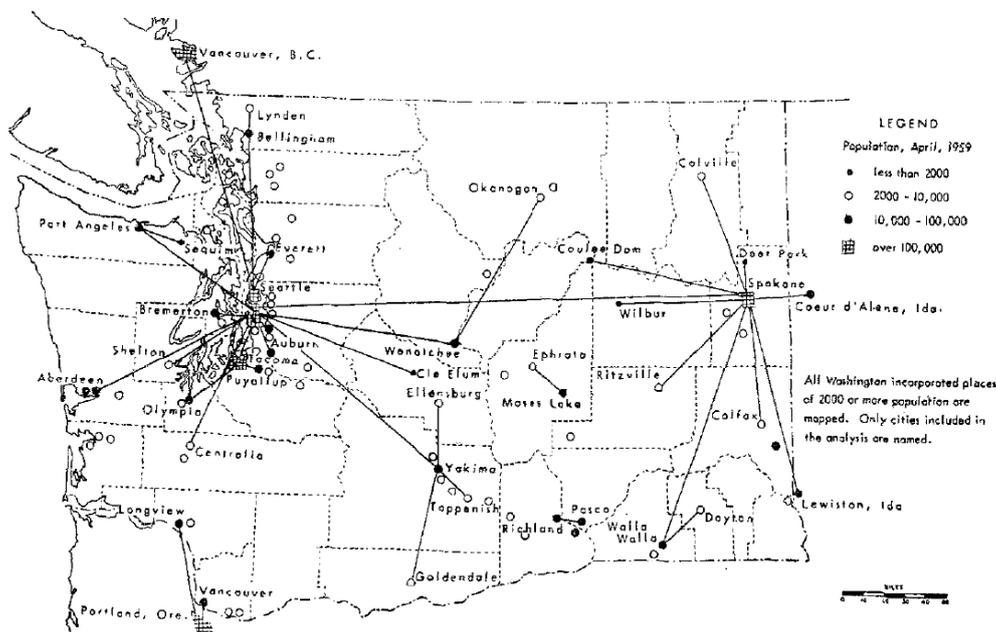
9.1.3 Cartographie de la structure nodale et régionalisation

La cartographie de la structure nodale, présentée sur la Figure 9-4, apporte deux nouvelles informations qui portent sur l'effectif de population des lieux, en mobilisant la *Taille* et sur une de leur caractéristique qualitative qui en découle à savoir leur attractivité, par une variation de *Forme*. Il est important de noter ici que J. D. Nyusten et M. F. Dacey ne sont pas tout à fait logiques. En effet, par rapport à la méthode qu'ils proposent, le choix naturel et logique d'utilisation de la taille (masse) des lieux aurait dû être la somme des flux dirigés vers un lieu (Dj) plutôt que la *Forme*.

Sur la carte de la Figure 9-4, les villes analysées sont assorties d'une étiquette mentionnant leur toponyme, celles qui n'en ont pas ont été exclues de l'analyse pour diverses raisons (Nyusten et

Dacey, 1961 : 40-42). Toutefois, toutes les villes de 2 000 habitants et plus ont été considérées, de même qu'elles sont représentées sur la carte.

Figure 9-4. Cartographie de la structure nodale des flux téléphoniques (Nyusten et Dacey, 1961)



Source : J. D. Nyusten et M. F. Dacey, 1961 : 39.

D'après les éléments portés en légende de la carte, le système étudié est formé de quatre catégories de villes distinguées en fonction de leur population et non de leur position dans le système de flux, ce qui aurait été pleinement logique. En effet, les figurés carrés ne sont pas liés au caractère de la ville (dominant/dominé) puisque la ville de Spokane (à l'est) en relie deux, ils sont plutôt utilisés pour représenter leurs différences de tailles, un critère associé de population. Le fait que les villes soient dominantes, dominées ou intermédiaires se déduit en réalité de la structure même du graphe, nous y reviendrons dans la section 9.2. Cependant, le fait que cette caractéristique ne soit ni exploitée ni explicitement cartographiée par des variables visuelles adéquates crée des ambiguïtés sur les relations. Dans le sud, par exemple, le lien de Pasco concerne deux cercles noirs de même qu'à l'est où le carré de Spokane relie celui de Seattle à l'est, ainsi que des cercles noirs (Dear Park au nord-est) et blancs (à Calfax, au sud).

On en conclue que, sur la Figure 9-4, l'alternance de teintes noires et blanches des cercles n'est pas liée au caractère dominant, dominé ou relais des villes, mais à la taille des villes, ce qui est pour le moins étrange.

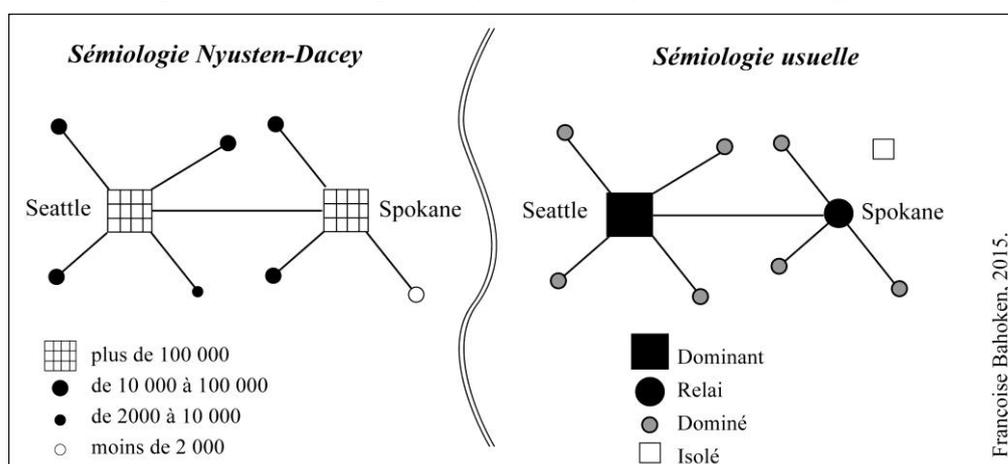
Plus généralement, deux éléments sont notables au niveau de la construction cartographique de la Figure 9-4. Les liens issus des « flux dominants » sont dirigés vers les différents pôles. Ils sont représentés par des traits droits, non valués et non orientés, à la manière d'oursins. **Les auteurs auraient donc fait le choix**, à l'instar d'une suggestion déjà mentionnée (Bertin, 1973 : 344, voir page 173) de dégrader l'information de départ (Fij) ← (Lij), de manière à ne cartographier que les liaisons (Lij) existantes résultantes des flux dominants ; **de représenter cette information quantitative sur des flux dominants d'une part, de manière qualitative et d'autre part, en focalisant l'attention sur les lieux.**

Sur le plan analytique, les auteurs indiquent d'ailleurs que les méthodes mobilisées visent à produire une régionalisation polarisée sur des villes qui « [...] se partageront un ensemble de villes en sous-

groupes qui spécifient une place centrale et sa hiérarchie subordonnée. ». En effet, les liens illustrant l'attractivité des points nodaux (les pôles) décrivent leur zone d'influence qui englobe un ensemble de villes de moindre importance (depuis lesquelles ils reçoivent des flux dominants). Ces liens décrivent la structure spatiale inhérente à chaque point nodal qui correspond aux plus grandes villes de l'état de Washington DC (ou des états voisins) ; les autres lieux sont des points de départ de l'arborescence (dominé) ou des relais (dominants et dominés).

On observe également des villes qui jouent un rôle de relais majeur sans pour autant être dominante : c'est le cas de Spokane pour Seattle. Pour cette raison, cette ville aurait dû être représentée sur la Figure 9-4 avec un symbole distinct de celui de Washington DC, par ailleurs l'unique pôle dominant de la zone d'étude. Seattle est, en effet, l'unique pôle nodal de cette ensemble et Spokane en est l'un des relais, certes le plus important, mais cela ne justifie pas qu'il ait une sémiologie identique au pôle. Raisonner ainsi est source d'ambiguïtés. Aussi proposons-nous de replacer cette lecture spécifique du caractère dominant ou relais des lieux dans la sémiologie usuelle. La Figure 9-5 compare la sémiologie de Nyusten-Dacey à l'approche classique de sémiologie cartographique des entités ponctuelles présentant un caractère qualitatif, l'objectif étant de distinguer la « capitale régionale », le pôle nodal de cette zone, des autres villes d'importance ou non de l'état.

Figure 9-5. Sémiologie de Nyusten-Dacey versus sémiologie usuelle



La partie gauche de la Figure 9-5 reprend un détail de la carte précédente (une sélection des relations de Seattle et de Spokane) pour illustrer la sémiologie de Nyusten-Dacey. L'objectif est en particulier de montrer que les auteurs mobilisent une *Forme* identique pour symboliser les pôles nodaux (Seattle) et intermédiaires (Spokane) – car ils raisonnent sur leurs masses (ici les effectifs de populations) alors qu'ils analysent l'attractivité des lieux ; des *Formes* variées pour figurer les lieux dominés. On observe en effet que des lieux satellites de Spokane sont représentés par des cercles noirs ou blancs de tailles variées et non par des cercles de taille proportionnelle en fonction du caractère correspondant (la légende indique que ces cercles représentent des données quantitatives discrètes, un effectif de population, ...).

La partie droite de la figure correspond à une version de la sémiologie usuelle permettant de représenter l'attractivité des lieux dominants, à l'aide de *Formes* variées : des carrés et de cercles. M. Baron et al. utilisent, par exemple, trois types de symboles circulaires : un petit cercle noir pour les lieux dominés, le même cercle noir entouré d'un liseret noir ou rouge, pour symboliser respectivement les pôles relais et dominants (Baron et al., 2008 : 5). Dans notre cas, le pôle dominant est dissocié des autres lieux, il est représenté par un carré noir de grande taille. On notera que les pôles universitaires, qui font l'objet d'une carte, sont également représentés par un carré par (Baron et al., 2008 : 13). Sur la

Figure 9-5, les autres lieux (hors pôles) sont symbolisés par un petit cercle noir ou gris selon qu'ils sont relayeurs ou dominés ; le lieu isolé est représenté par un carré blanc.

Cette proposition correspond en fait au croisement des différentes caractéristiques des lieux envisageables, selon qu'ils sont dominants et / ou (non) dominés, comme illustré ci-dessous.

Figure 9-6. Les quatre formes corrigées de la sémiologie cartographique de Nyusten-Dacey

		<i>Lieux dominés</i>	
		○ oui	■ non
<i>Lieux dominants</i>	oui ■	dominant et dominé ● <i>lieu intermédiaire (relai)</i>	dominant et non dominé ■ <i>pôle dominant (nodal)</i>
	non □	non dominant et dominé ● <i>lieu satellite</i>	non dominant et non dominé □ <i>lieu isolé</i>

Françoise Bahoken, 2015.

D'après la Figure 9-6, l'approche Nyusten-Dacey correspond à quatre possibilités associées à autant de symboles. Ces possibilités peuvent être regroupées en deux catégories graphiques, c'est-à-dire associées à deux *Formes* sur le plan sémiologique : celle des lieux dominants que nous avons symbolisée par un carré et celle des lieux non dominants ou dominés, figurée par un cercle. C'est parce que ces catégories ne sont pas toujours imperméables que les symboles correspondants peuvent être combinés entre eux (et / ou en association avec la *Couleur*), correspondant à quatre cas de « figures » *a priori* envisageables :

- le cas des lieux qui sont à la fois dominants et dominés, qui caractérisent les lieux intermédiaires ;
- le cas des lieux dominants et non dominés, correspondant aux pôles nodaux ;
- le cas des lieux non dominants et non dominés qui défini les lieux isolés ;
- le cas des lieux non dominants et dominés qui caractérisent les lieux satellites.

Les cartes des *flux dominants* (issues de la Figure 9-5, moyennant une sémiologie adaptée) traduisent ainsi indéniablement la volonté des auteurs de focaliser l'analyse davantage sur les lieux que sur les flux, en utilisant une variation de *Forme* pour distinguer les lieux selon leur taille, leur attractivité. Sur la Figure 9-5, on voit bien que la sémiologie des lieux prime sur celle des liens et cela, quelle que soit la méthode utilisée. La représentation en oursins qui en découle renforçant en outre l'effet d'attractivité de ces lieux d'intérêts. Cela s'explique aussi par le cadre de référence des auteurs qui est celui de la théorie des graphes, où la représentation du graphe d'adjacence ou de connexité prime sur le reste (la valeur des liens et la localisation des lieux).

En outre, une spécificité de la méthode qu'il convient de mentionner porte non seulement sur la sélection drastique des flux obtenue par la méthode, mais sur l'identification de régions dites nodales indépendantes formées par un ensemble de liens définissant un sous-graphe connexe. Ces régions ou sous-régions, indépendantes ou disjointes les unes des autres, sont formées d'un ensemble d'entités partageant une caractéristique d'appartenance territoriale commune (voir section 7.3, page 306) : celle d'être sous l'influence d'un même pôle.

De ce point de vue, la méthode Nyusten-Dacey procure trois familles de résultats cartographiques :

- une cartographie des seuls flux de rang 1, suite à la restriction locale de la matrice (Fij) aux seuls *flux majeurs*²¹². L'approche locale permet de conserver l'ensemble des lieux du système ;
- une analyse de la distribution spatiale des (i) lieux interprétée en termes *d'attractivité* (Dj), par le recours à un critère de dominance fondé sur l'analyse des valeurs de flux $\text{Max}_i(\text{Fij})$;
- une *régionalisation* de la zone d'étude, en régions ou sous-régions nodales (Pj), correspondant à une partition de la matrice (Fij), selon le critère de dominance des lieux.

En généralisant le propos, on observe que l'application d'un critère de sélection fondé sur la localisation des lieux intéresse par conséquent non seulement la famille de méthodes de cartographie de flux liées à la restriction des matrices, mais également celles qui sont liées à leur partitionnement dans l'objectif de mettre en évidence le découpage de la zone d'études en structures polarisées autonomes. Elle « [...] doit être comprise comme une méthode révélant des structures et des hiérarchies régionales, mais non des régions ; elle devra être complétée par une procédure de partition spatiale » (Haggett 1973, cité par Dumolard, 1975 : 07).

L'analyse combinant approches graphiques et analytiques, la méthode Nyusten-Dacey correspond à l'une des familles les plus importantes de l'analyse cartographique de flux, de même qu'elle est la plus riche sur le plan sémiologique. Pour le démontrer, nous allons examiner ses différents cas d'application, en mobilisant la matrice (Fij) de l'Isère ; la validation empirique à l'échelle mondiale sera réalisée sur la matrice des flux commerciaux internationaux de l'OMC (2007), mais elle possède quelques défauts d'ordre sémiologique que nous pouvons corriger.

9.2 Proposition de reconstruction cartographique de la méthode Nyusten-Dacey (1961)

La méthode présente une particularité liée au fait qu'elle s'inscrive dans la théorie des graphes. En effet, en étant fondée sur un raisonnement géométrique purement topologique, la représentation des *flux dominants* est réalisée sous la forme d'un graphe de connexité où les liens ont une largeur unique. Cette cartographie de flux ne tient pas compte de leur valeur (Fij), mais seulement de l'existence d'un lien qui en découle (Fij). Elle apparaît alors focalisée sur les lieux, en particulier sur l'une de leurs caractéristiques qualitatives conduisant à une analyse de leur hiérarchie.

En effet, la cartographie de *flux dominants* (voir Figure 9-1, page 351) représente des liens (Lij) assortis de symboles aux formes géométriques variées. Quelles en sont les raisons ?

On notera au passage que la carte de P. Mariot (Figure I-1-1, page 13) pourrait d'ailleurs être construite selon ce principe des *flux dominants*.

Nous proposons d'analyser le processus de construction d'une carte de *flux dominants*, sous un double angle cartographique et sémiologique. Cela nous conduit à l'examen des cinq principales étapes suivantes.

²¹² C'est cette méthode de réduction de la matrice de flux a été appliquée par N. Cattani (2004), dont il a été question dans le chapitre 7.

- **Etape 1 : Cartographie des flux majeurs** : correspondant à l'application d'un premier critère de sélection des flux fondé sur leur valeur ($F_{ij} \leftarrow \text{Max}_i(F_{ij})$) ;
- **Etape 2 : Cartographie de l'attractivité des lieux** devant définir leur caractère dominant ou non elle correspondant au second critère mobilisé (D_j), la variable (D_j) correspondant à la somme marginale en colonnes ;
- **Etape 3 : Cartographie des lieux et des liens dominants** : consiste en la combinaison des premiers et seconds critères : ($F_{ij} \leftarrow \text{Max}_i(F_{ij})$) et ($D_j > (D_i)$) ;
- **Etape 4 : Cartographie des régions nodales**, c'est-à-dire de l'aire d'influence (A_{ij}) des pôles nodaux ;
- **Etape 5 : Cartographie des sous-composantes (sous-régions) nodales**, réalisée en complément de l'étape précédente.

Nous allons montrer que chacune de ces cinq étapes correspond à une logique de raisonnement particulière mobilisant des implantations spatiales et les principes de sémiologie cartographique qui leur sont associées. Sur le plan théorique, nous verrons que cette méthode privilégie une *logique de lieux-liens* plutôt que la *logique de flux* classique, dans le cadre d'un raisonnement en partie qualitatif (recourant à la *Forme* et à la *Couleur*) plutôt qu'exclusivement quantitatif (ne mobilisant que la *Taille*).

9.2.1 Etape 1 : cartographie des flux majeurs $F_{im} = \text{Max}_i(F_{ij})$

La première étape de l'analyse consiste à sélectionner un seul flux ($k=1$) et le plus grand ou celui de rang 1 : $\text{Max}_i(F_{ij})$ émis par les lieux. Sur le plan formel, la méthode consiste à transformer la matrice (F_{ij}) en matrice (ou graphe) de type *k-core*, où chaque sommet est relié à au moins un autre mais en tenant compte d'une condition supplémentaire correspondant à la valeur maximale, au flux de rang 1 ($r=1$) ou majeur (voir [9.1]).

La partie droite de la Figure 9-7 illustre la sélection des seuls *flux majeurs* de la matrice exemple. Ils sont portés en grisé et correspondent aux flux émis de rang 1.

Figure 9-7. Illustration de la sélection de *flux majeurs* de (F_{ij})

Valeurs disponibles					Rang des flux				
Fij	A	B	C	D	Fij	A	B	C	D
A	-	8	5	1	A	-	1	2	3
B	6	-	3	2	B	1	-	2	3
C	10	4	-	1	C	1	2	-	3
D	1	2	3	-	D	3	2	1	-
Dj	17	14	11	4	Dj	17	14	11	4
r	1	2	3	4	r	1	2	3	4

Les valeurs de flux majeur sont représentées en grisé.
r : rang croissant par rapport aux lignes

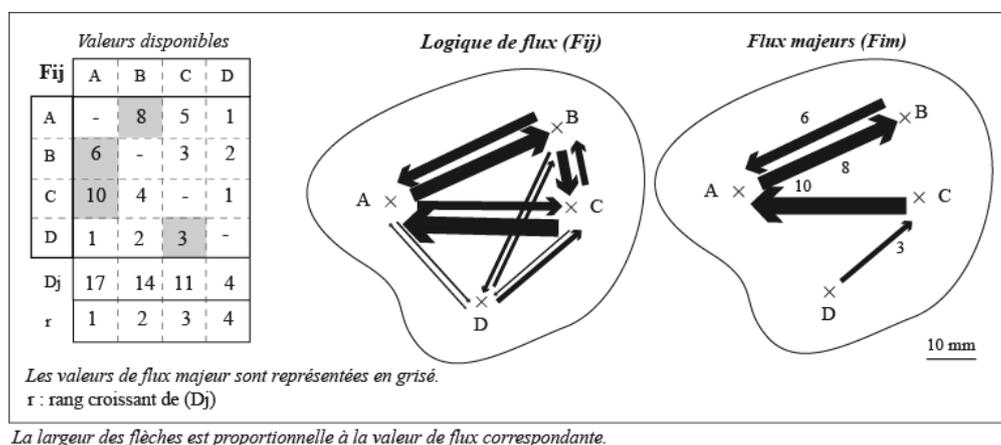
L'opération consiste en une transformation des valeurs quantitatives discrètes en valeurs ordinales, de rang qui sont interprétés du point de vue des seuls lieux d'origine (nous reviendrons sur la prise en compte des lieux de destination dans la section 9.3 (voir *infra*)).

Cependant, cette transformation ordinale de la matrice peut poser problème. D'après P. Haggett, l'approche soulève l'un des premiers problèmes des statistiques non paramétriques, lié à « [...] l'abandon d'une partie de l'information qu'elles contenaient » (Haggett, 1973 : 320). Pour y remédier, de nombreux auteurs ont conservé la structure quantitative de la matrice (partie gauche de la Figure 9-7) plutôt que la version ordinale (partie droite de cette même figure).

Cela conduit à représenter des flux et non, comme le suggèrent J. D. Nyusten et M. F. Dacey, les liens correspondant aux *flux majeurs*. La différence entre ces deux possibilités est présentée sur la Figure 9-8.

Ce choix de conserver les valeurs de flux (de cartographier des *flux majeurs*) à trois conséquences sur leur cartographie. Il implique de représenter i) les valeurs de ces flux de manière usuelle (en faisant varier leur *largeur*, en application de la *Taille*) ; ii) de même que leur orientation, en introduisant la flèche et éventuellement, iii) les valeurs de flux en les discriminant mieux, par une action sur la teinte des flux. Cette possibilité est illustrée sur la partie droite de la Figure 9-8.

Figure 9-8. Illustration de la cartographie des *flux majeurs* de (Fij)



La cartographie des *flux majeurs* correspond, en fait, à la *logique de flux* classique (voir chapitre 3 et chapitre 6). Elle est, de ce fait, assimilable à une variante de solutions conduisant à cartographier des matrices creuses (Chapitre 8), des matrices préalablement réduites par l'application d'un critère de sélection unique, ici : $(Fij) \leftarrow \text{Max}_i(Fij)$. Autrement dit, la cartographie des *flux majeurs* n'apporte rien de nouveau, si ce n'est que le critère mobilisé agit localement (par ligne) et non pas globalement.

Cette cartographie des seuls *flux majeurs* s'inscrit dans une *logique de flux* classique, menant à la cartographie de données quantitatives (Fij) qui portent sur les liens (Lij), elle ne fait pas nécessairement appel aux lieux et à leurs caractéristiques. Le passage à la cartographie des *flux dominants* opère un double changement, dans la mesure où le critère « *dominant* » est de type qualitatif (« *A domine B* ») et interprété du point de vue des lieux A et B, et non plus des liens. C'est pourquoi la seconde étape porte sur le critère de domination des lieux.

9.2.2 Etape 2 : d'une *logique de flux* à une *logique de liens-lieux*

La seconde étape de l'analyse concerne la caractérisation des lieux, selon qu'ils sont attractifs ou non, devant conduire à l'identification – à l'étape suivante – des *flux dominants*. Elle s'appuie sur la seconde condition introduite par les auteurs, qui mobilise la somme marginale en colonne (Dj) pour comparer deux lieux. Par exemple, la comparaison de $(Dj)_A$ et de $(Dj)_B$ sur la Figure 9-8 montre l'attractivité supérieure du premier sur le second : $(Dj)_A > (Dj)_B$ ou A plus attractif que B.

Plusieurs éléments sont notables à ce stade, tant sur le plan de la construction cartographique qu'au niveau analytique.

Sur le plan cartographique, le recours à l'attractivité (Dj) des lieux mobilise une implantation ponctuelle, associé à un symbole graphique : un cercle, par exemple. Les auteurs ont d'ailleurs proposé une illustration de cette hiérarchisation qualitative des lieux, considérant la notion de hiérarchie comme l'une des propriétés fondatrice de la notion de « *flux le plus important* ».

« *L'une des propriétés stipule qu'une ville est 'indépendante' si son flux majeur provient d'une ville plus petite [...] pour les identifier, une mesure de leur taille est requise. La taille peut être une donnée exogène, comme par exemple, la population de chaque ville ; ou intrinsèque à la matrice représentée, par exemple, le volume total de messages émis ou reçu par l'ensemble des villes de la région concernée [...] Une ville dominée est une ville dont le flux majeur est destiné vers une ville plus importante* » (Nyusten et Dacey, 1961 : 34).

Les lieux peuvent ainsi être caractérisés en deux grandes familles, selon qu'ils « *dominant* » ou qu'ils sont « *satellite* ». La donnée (Dj) étant quantitative, mais son interprétation qualitative (dominant / satellite), elle ne mobilise la *Taille* pour représenter le fait qu'un lieu A est plus gros qu'un lieu B, mais la *Forme* de sorte que la carte indique si A est *dominant* ou non, autrement dit si A *domine* B est satellite). Le raisonnement aurait, tout aussi bien pu mobiliser *l'Orientation* – mais ce n'est pas le choix des auteurs – et la *Couleur*, mais nous souhaitons conserver cette dernière variable visuelle pour les étapes 4 et 5 (voir *infra*).

Il est important de noter que ce raisonnement focalisé sur la forme des symboles qui décrivent les lieux, pour les qualifier, entraîne **le basculement du raisonnement quantitatif continu mobilisé jusqu'ici pour cartographier des flux associée à la Taille vers un raisonnement qualitatif discret mobilisant la Forme et, plus loin, la Couleur. De même que la logique du raisonnement varie, en passant de la logique de flux ordinaire à une logique de lieux et de liens.**

Sur le plan analytique, la méthode mobilise donc désormais des données qualitatives associées à chaque lieu, ce qui modifie la signification de la carte. Effectivement, ce qui importe ici c'est le fait qu'un lieu A domine un lieu B, que la relation soit binaire (dominant / non dominant) et orientée, plutôt qu'un lieu A soit plus gros ou plus grand qu'un lieu B.

Ce basculement sémantique n'est pas sans conséquences sur le plan sémiologique. Il entraîne d'une part, la mobilisation d'une implantation ponctuelle (pour les lieux) et d'autre part, la suppression de la valeur des liens.

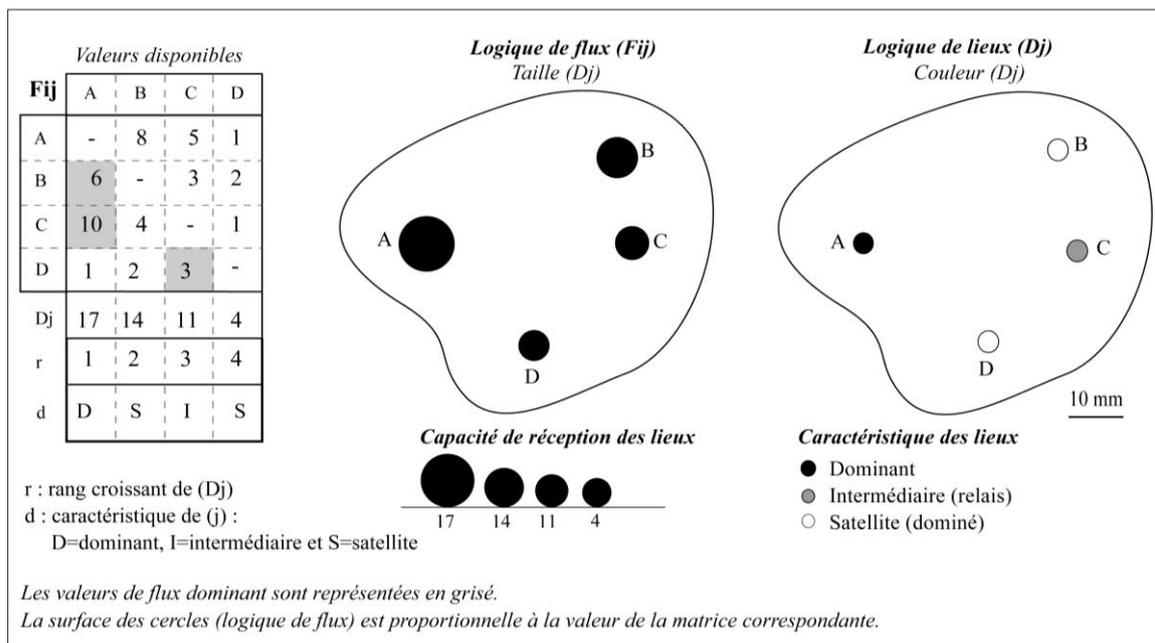
Dans le premier cas, il convient de choisir une forme ponctuelle, ici un cercle, ce qui entraînera la mobilisation des variables visuelles associées à la représentation de données ponctuelles, présentant des caractéristiques qualitatives. Ces caractéristiques sont portées sur le tableau, à la ligne d.

La cartographie de l'attractivité d des lieux de destination selon qu'ils sont dominants ou satellites, est présentée sur la Figure 9-9.

La partie gauche de cette même figure illustre la mise en oeuvre du cas classique de représentation de l'attractivité des lieux appréhendée de manière quantitative, par la *Taille*. La partie de droite décrit l'approche Nyusten-Dacey qui mobilise en fait la teinte : le symbole du lieu teinté de noir correspond au lieu de rang 1, les autres lieux sont symbolisés par un cercle blanc ou grisé si le lieu est à la fois dominant et dominé.

Ces deux acceptions du même processus d'attractivité à représenter signalent le basculement du raisonnement d'une *logique quantitative* à une *logique qualitative*.

Figure 9-9. Illustration de la cartographie de l'attractivité des lieux (Dj)

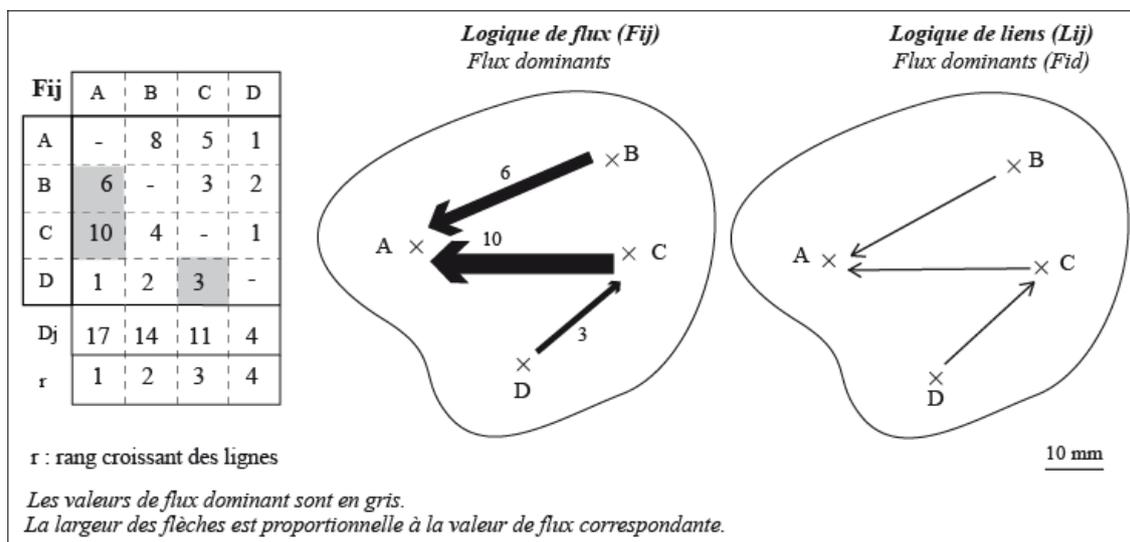


On observe en effet les conséquences sémiologiques de l'effet du basculement d'une *logique de flux* (Fij) *quantitative* à une *logique de lieux* (Dj) *qualitative*, qui mobilisent respectivement la *Taille* et la *Forme* associée à la *Couleur*, ici dans l'esthétique du noir-et-blanc (nous souhaitons, en effet, réserver l'usage d'une gamme colorée pour les deux dernières étapes).

La seconde modification sémiologique, apportée à cette étape du processus, est une conséquence de la première. La focalisation de l'analyse sur les valeurs des lieux supprime celles des flux qui ne seront dès lors pas représentés. Seuls les liens (Lij) traduisant l'existence de flux dominants seront cartographiés.

La Figure 9-10 illustre cette modification conceptuelle conduisant à la cartographie de liens non valués et orientés, en remplacement des flèches. Elle met en évidence la complémentarité de l'analyse liens-lieux où le lieu de rang 1 identifié sur la Figure 9-9 et représenté par un cercle noir, est aussi celui qui polarise les flux sur la Figure 9-10. Il s'agit du lieu A.

Figure 9-10. Illustration de la cartographie des liens (Lij) issus des flux dominants (Fid)



Si l'on revient sur la Figure 9-4 (page 354), on observe que J. D. Nyusten et M. F. Dacey ont bien représenté les lieux les plus attractifs (de rang 1 des sous-systèmes) par des symboles dont la teinte varie, mais qu'ils ont également mobilisé une variation de forme géométrique. Les auteurs ont ainsi utilisé la *Forme* pour discriminer les types de lieux (des cercles et des carrés), associée à une modification de leur teinte (les cercles sont blancs ou noirs), autrement dit en mobilisant la *Forme* en association avec la *Couleur*.

L'introduction de variations de *Forme* concerne l'étape 3 suivante, focalisée sur la combinaison des deux sous-parties de cette seconde étape sur la même carte (cela correspond à la fusion des parties droites des cartes de la Figure 9-9 et de la Figure 9-10 de la page précédente).

9.2.3 Etape 3 : cartographie des « flux dominants »

La troisième étape porte sur la représentation de la combinaison de l'attractivité des lieux (par des symboles présentant la même forme, mais étant de tailles variables), dans le système de relations issu de la sélection des *flux majeurs* (par des liens non valués et non orientés). Elle définit la cartographie des *flux dominants* proprement dite, une représentation des flux selon une *logique de lieux* ou du point de vue des lieux.

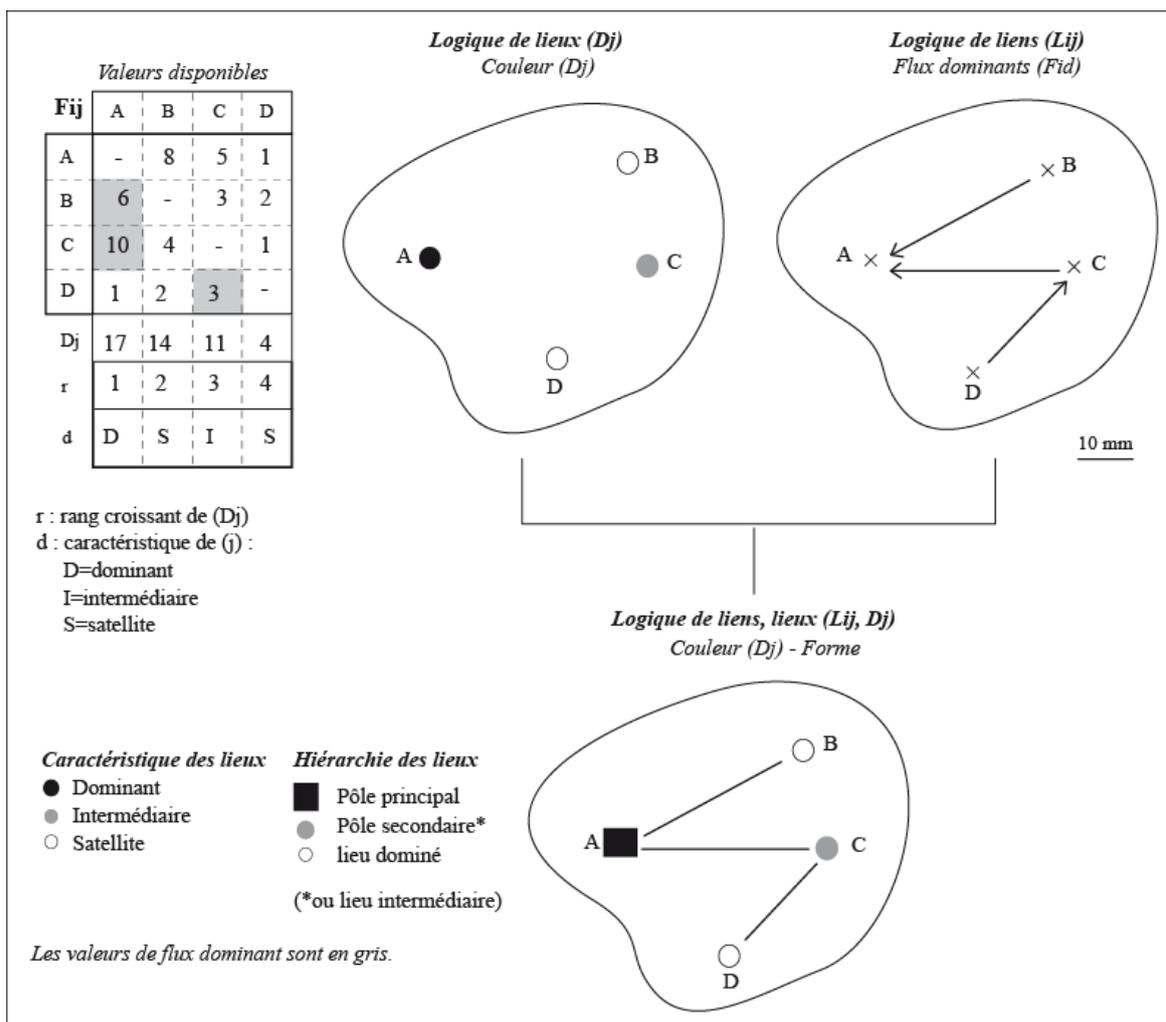
Cette combinaison *liens-lieux* pose toutefois de nouveaux²¹³ problèmes de sémiologie cartographique. Les difficultés sont liées à l'analyse de la carte, passant par la gestion des relations transitives considérées, d'après les auteurs, comme seconde caractéristique de la méthode qu'ils définissent de la manière suivante. « Cette propriété implique que si une ville *a* est subordonnée à une autre ville *b* et que *b* est subordonnée à *c*, alors *a* est subordonnée à *c* » (Nyusten et Dacey, 1961 : 34). De même, il est nécessaire de gérer, d'une part, les positions de villes ex-æquo ($D_j = D_i$) et, d'autre part, celles des lieux intermédiaires qui ne sont ni dominants ni dominés correspondant à deux problèmes distincts, mais que les auteurs considèrent comme une troisième propriété stipulant qu'aucune ville ne peut être dominée par l'une de ses subordonnées. Pour y parvenir, il convient d'introduire des variations de *Forme*, en introduisant des symboles différents pour catégoriser les lieux, en plus de leur hiérarchisation qualitative déjà introduite par la *Taille*. On mobilise alors deux autres propriétés associées à la caractéristique des lieux : la *Couleur* et la *Forme*, tel que illustré sur la figure suivante

La cartographie des *flux dominants* répond strictement à la double condition des auteurs, en combinant la partie droite de la Figure 9-9 et de la Figure 9-10. Cette combinaison est présentée sur la Figure 9-11, introduisant la *Forme* : elle distingue les différentes catégories de lieux de sorte que la carte montre la situation des très grands pôles (lieux dominants) qui structurent le système des flux à un niveau régional plus que local, ce dernier correspondant aux lieux satellites. C'est pourquoi, sur le plan sémiologique, ces lieux sont représentés par un carré (par exemple), de la même manière que l'on distingue les capitales nationales des autres villes d'un Etat dans la pratique usuelle.

L'introduction de la *Forme* conduit à visualisation de la structure hiérarchique des lieux en trois catégories : un pôle principal ; un pôle secondaire, pour illustrer la « *relation de transitivité* » évoquée par les auteurs et un lieu satellite (ou lieu dominé). A chacune de ces catégories correspond ainsi un symbole particulier dont la taille est fonction de leur qualité. C'est pourquoi il est possible de supprimer l'orientation des liens (la flèche) issus des flux dominants, dans la *logique liens-lieux*, puisque la hiérarchie des lieux indique le sens de la domination.

²¹³ L'effet de taille associé à l'effet de convergence des flux vers les zones de plus forte attractivité sont, tous deux, considérés ici comme des problèmes anciens.

Figure 9-11. Illustration de la hiérarchie des lieux du système de flux dominants (Fid)



Cette typologie est cependant souvent insuffisante, car elle ne permet pas de caractériser les lieux ayant une position intermédiaire dans le système de flux dominants si ce dernier présente une arborescence compliquée, ou plusieurs dominations successives. C'est pourquoi différents auteurs en ont proposé des typologies en trois voire quatre classes qui permettent d'affiner la caractérisation des lieux. N. Cattan (2001) propose, par exemple, de définir de la manière suivante les lieux, selon qu'ils sont *dominants*, *intermédiaires*, *dominés* ou *satellites* et *indépendants* :

- un lieu est *dominant* s'il reçoit un *flux majeur* en provenance de lieu(x) de plus petite taille. Il s'agit de lieux qui dominent les autres, sans toutefois être dominés ;
- un lieu est *intermédiaire* s'il est dominé par un pôle, ou un lieu de taille plus importante et qu'il domine, par ailleurs, au moins un autre lieu de taille inférieure ;
- un lieu est *dominé* s'il ne reçoit aucun *flux majeur* en provenance d'autres lieux ;
- un lieu est *indépendant* s'il ne reçoit aucun *flux majeurs* d'autres lieux et si son propre *flux majeur* est émis vers un lieu de plus petite taille.

Les lieux dominés sont, à l'inverse des pôles nodaux, ceux qui sont placés dans l'aire d'influence de ces derniers ; ils appartiennent à la même sous-composante connexe sans pourtant occuper la même position dans le système de relation, ce qui signifie que le graphe associé à la cartographie de flux

dominants peut être interprété en termes de régionalisation, comme nous l'avons vu dans la section 7.3 (page 306).

9.2.4 Etapes 4 et 5 : régions et sous-régions polarisées

La quatrième étape est celle de la régionalisation de l'espace pouvant être déduite de l'analyse des *flux dominants*. En effet, nous exploitons ici le fait que la méthode de Nyusten-Dacey « [...] *doit être comprise comme une méthode révélant des structures et des hiérarchies régionales, mais non des régions ; elle devra être complétée par une procédure de partition spatiale* » (Dumolard, 1975 : 107). Pour ce faire, le raisonnement que nous allons mener est de nature qualitative, il consiste à définir l'aire d'influence des pôles correspondant à une région nodale.

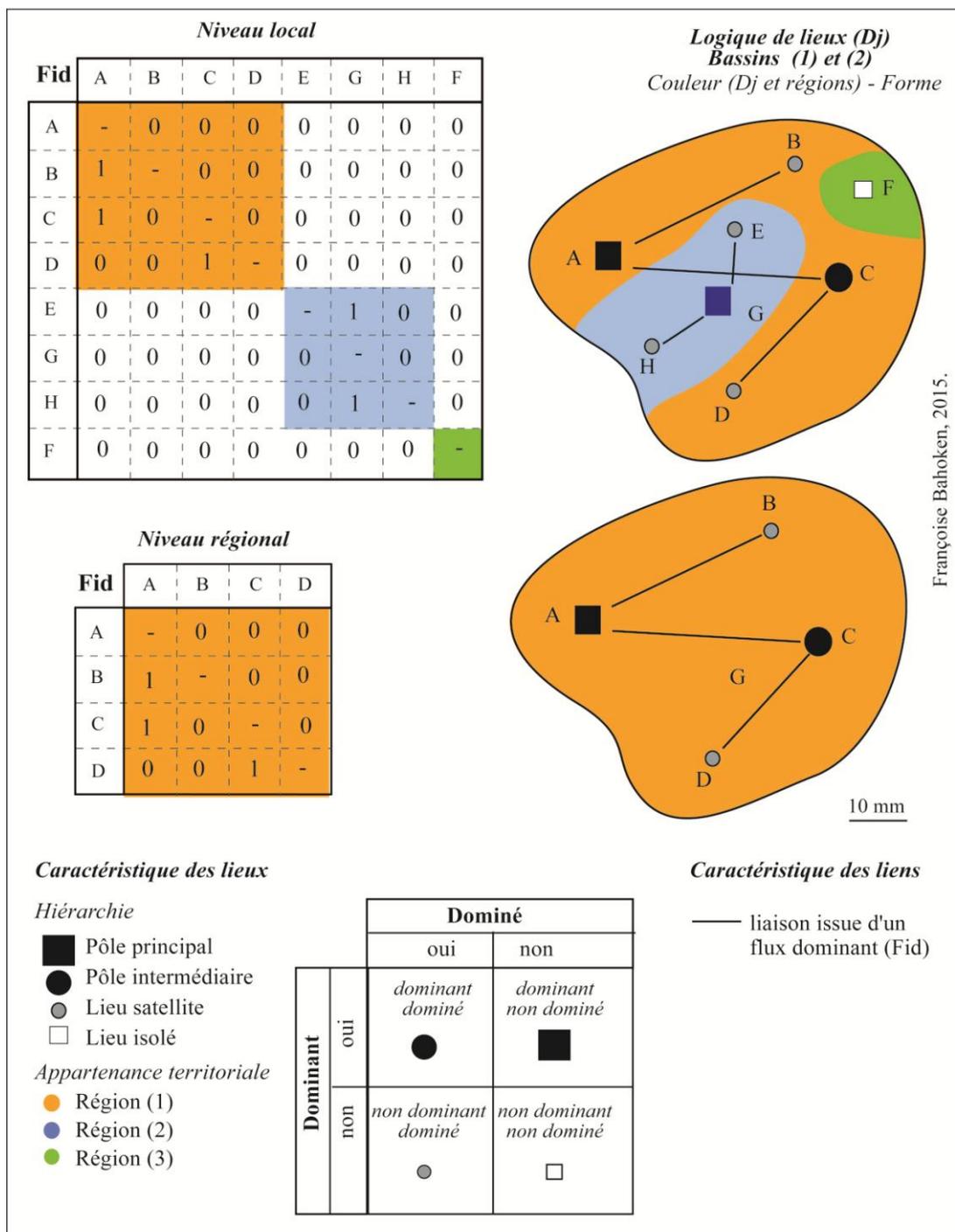
Les lieux non nodaux (non dominants), satellites ou intermédiaires font, en effet, partie de la zone d'influence (ou de domination) d'un (ou de plusieurs) pôles nodaux et leur caractérisation dépendant de leur position dans le graphe associé, examinée à l'étape précédente.

Le recours à la notion d'aire d'influence englobant un ensemble de lieux introduit un nouvel objet géographique : la *région* (définissant l'aire) et, *à fortiori*, un nouveau mode d'implantation spatiale : la surface, ainsi que les variables visuelles qui sont associées à sa représentation, principalement la *Couleur* (ou le *Grain*) et la *Valeur*, si on souhaite ajouter des sous-régions.

Pour illustrer cette quatrième étape de régionalisation issue de l'analyse des *flux dominants*, nous allons mobiliser la matrice exemple local de rang 8, dont la construction a été présentée sur la Figure 7-16 (page 307). La procédure de création des aires d'influence étant, en fait, une variante de la construction d'une matrice d'appartenance régionale (A_{ij}) dont les régions seraient formées de villes satellites et/ou intermédiaires dominées par un pôle, dans une structure en arbre hiérarchique.

Toutefois, pour les besoins de la démonstration, cette matrice fictive fait l'objet d'un réordonnement de sorte que puissent apparaître des régions qui seraient issue des trois étapes précédentes. La première étape de la régionalisation, d'identification des bassins principaux est présentée sur la Figure 9-12. Elle mobilise la sémiologie cartographique de Nyusten-Dacey modifiée, telle que présentée sur la Figure 9-5, page 355.

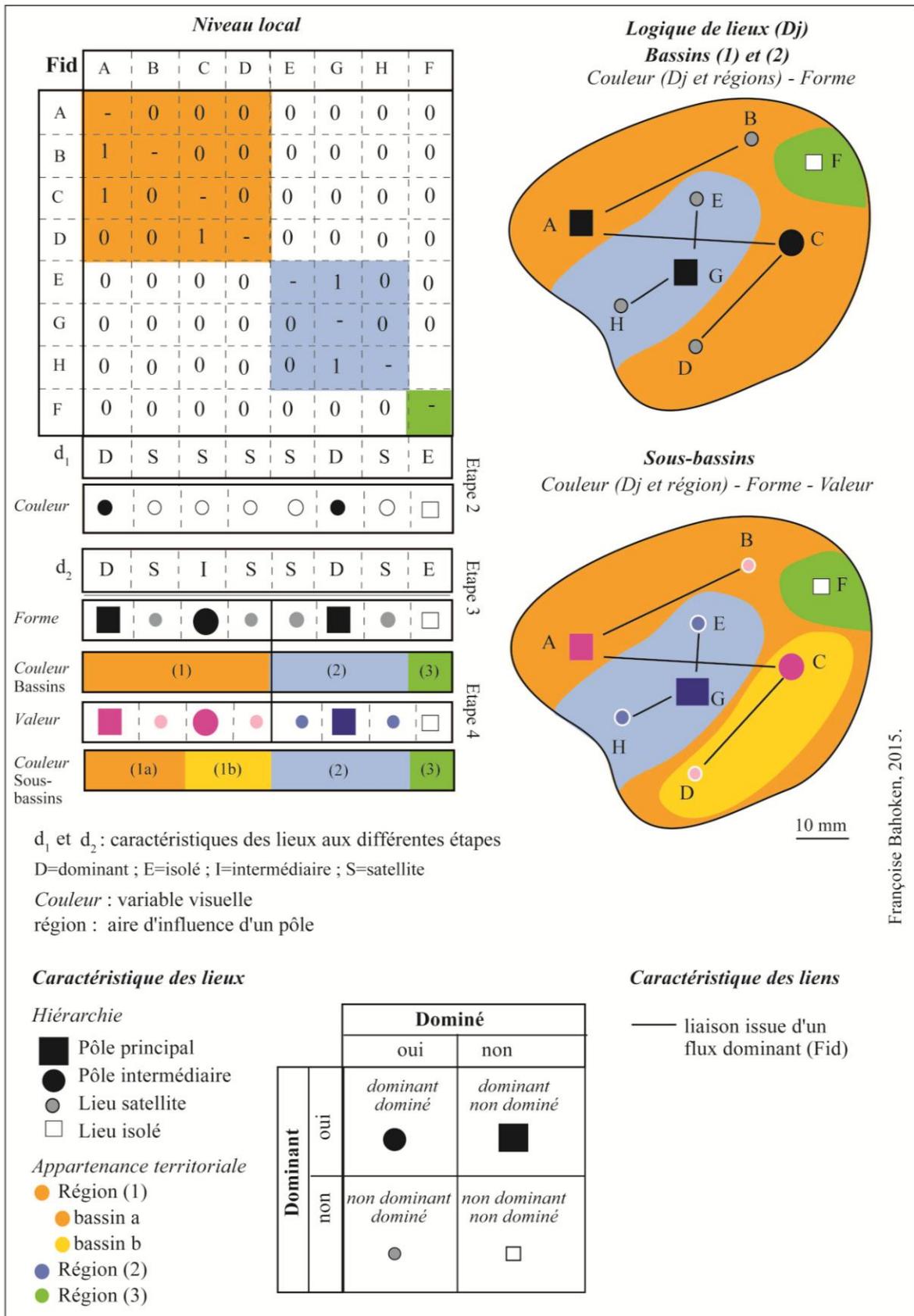
Figure 9-12. Illustration de la régionalisation issue du système de flux (Fij) dominants (Fid)



Ainsi, en supposant que notre matrice exemple corresponde à la région (1), formée des lieux A, B, C et D, nous lui appliquons une teinte orange pour la distinguer des autres. Nous pouvons ensuite mettre en évidence deux autres sous-composantes de cette matrice : la première correspond à la région (2), de couleur bleue et formée de la triade E, G, H et la seconde entité – ou région (3) – correspond à un singleton, formé d'un lieu F isolé. Ces différents lieux sont symbolisés dans le formalisme sémiologique présenté sur la Figure 9-6, page 356.

Le second temps de cette étape de régionalisation consiste à cartographier les sous-bassins des régions principales afin de hiérarchiser les lieux concernés. Elle est présentée sur la Figure 9-13.

Figure 9-13. Illustration des sous-bassins régionaux issus du système de flux (Fij) dominants (Fid)



Françoise Bahoken, 2015.

La Figure 9-13 mobilise en outre la *Valeur* pour caractériser la position des différents lieux intermédiaires et satellites appartenant au même bassin, ce qui soulève une nouvelle série de problèmes d'ordre analytique et sémiologique. Elle reconstruit ainsi l'ensemble de la procédure de cartographie de *flux dominants* selon la méthode Nyusten-Dacey (1961), en incluant les étapes de régionalisation et de sous-régionalisation pour la région (1), pourtant non représentées par les auteurs.

Sur le plan analytique, il convient de caractériser les sous-bassins en définissant une méthode de régionalisation (de formation des régions et sous-régions). Deux possibilités sont envisageables selon que l'approche est descendante ou bien ascendante. Dans le premier cas, probablement le plus facile à mettre en œuvre, l'approche consiste à décomposer la formation en sous-bassins en identifiant les lieux placés sous la domination d'un pôle par ordre d'importance décroissance : on identifie ainsi d'abord les lieux dominants puis les lieux intermédiaires et enfin les lieux satellites ou dominés. Dans le second cas, la vision est ascendante : elle consiste à partir du niveau local, de l'échelon territorial le plus fin, puis d'identifier progressivement les pôles intermédiaires et enfin, le pôle nodal. Comme nous l'avons déjà mentionné dans le chapitre 7, plusieurs possibilités sont, en effet, envisageables pour produire une régionalisation spatiale du point de vue des lieux.

Sur le plan sémiologique et quelle que soit l'approche analytique mobilisée, la méthode introduit *in fine* la *Valeur*, en association avec la *Forme-Couleur*, afin de distinguer la hiérarchie des lieux en fonction de leur appartenance régionale, à l'aire d'influence d'un même pôle.

La Figure 9-13 met en évidence la richesse de la méthode Nyusten-Dacey qui mobilise :

- les trois principaux modes d'implantations spatiales : des points, des lignes et des polygones ;
- les cinq principales variables visuelles associées à une cartographie *liens-lieux* des flux, que sont :
 - la *Localisation* des lieux, pour le semis de villes ;
 - la *Taille*, pour cartographier les *flux majeurs* (non représentés sur la Figure 9-13) ;
 - la *Forme*, pour hiérarchiser les lieux ;
 - la *Couleur* mobilisée deux fois, mais dans des styles différents en fonction de l'étape du raisonnement (en noir et blanc à l'étape 2 et en couleur à l'étape 4, pour identifier les régions) ;
 - la *Valeur*, pour caractériser les sous-bassins régionaux.

Illustrons maintenant le fonctionnement de cette méthode, en appliquant les cinq étapes présentées au cas de l'Isère (Fij), avant de les valider sur une matrice de flux internationaux.

9.2.5 Cartographie des flux dominants : l'exemple des navettes de l'Isère (Fij)

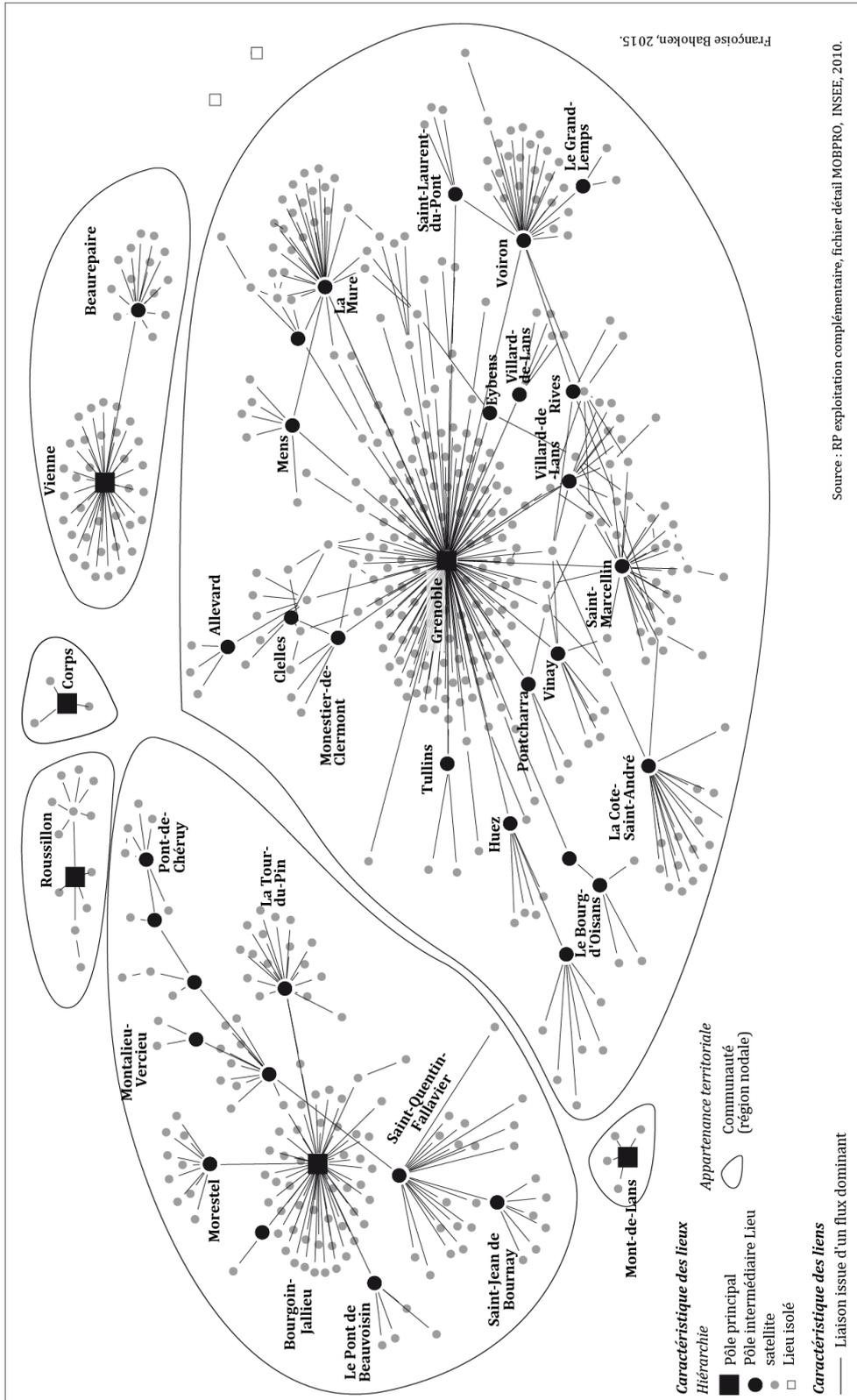
Nous avons vu que l'un des principaux intérêts de la méthode des *flux dominants* réside dans la mise en évidence d'une hiérarchie structurelle issue des flux. Sur le plan cartographique, cela passe par une représentation de l'attractivité des lieux qui en découle, en mobilisant la *Forme* et la *Couleur*. La méthode étant issue de la théorie des graphes, nous faisons le choix de présenter le graphe des relations nodales, obtenu à l'étape 3 précitée, puis la régionalisation que celui-ci procure, enfin la carte finale résultante de l'Etape 5. Nous allons ainsi au bout de la proposition de Nyusten-Dacey en proposant une cartographie de la régionalisation en sous-bassins issue de l'application de la méthode des *flux dominants*.

9.2.5.1 Représentation graphique des relations nodales

La Figure 9-14 présente le graphe des *flux dominants* de la matrice (Fij) de l'Isère obtenu à l'étape 3 précitée. Le placement des lieux résulte de l'application de la méthode de placement des sommets de

T. M. J. Fruchtermann et E. M. Reingold (1991), et la sémiologie est identique à celle de la Figure 9-5, (page 355). La matrice de l'Isère étant complexe, un certain nombre de lieux sont dominés par deux pôles sans pour autant être nécessairement dans une position intermédiaire. Afin de distinguer ces pôles *ex-aequo* au regard du lieu dominé en question, une règle arbitraire de rattachement de ce dernier à la plus grande sous-communauté a été définie.

Figure 9-14. Graphe des relations nodales de la matrice (Fij) de l'Isère



Le graphe de l'Isère met en ainsi évidence 6 régions principales (Bourgoin-Jallieu, Corps, Grenoble, Mont-de-Lans, Roussillon et Vienne) qui forment des sous-composantes connexes (ou des sous-régions) du graphe d'ensemble. Ces sous-composantes sont organisées autour de communes dominantes et qui le sont parce qu'elles fournissent des emplois à des communes dont la fonction est davantage résidentielle.

De fait, les sous-composantes du graphe peuvent être interprétées comme autant d'aires de recrutement des communes nodales, comme autant de bassins d'emploi, une hypothèse qu'il convient de vérifier, nous y reviendrons.

La commune de Grenoble est celle qui présente le maximum de communes placées sous son influence (près de 54% des communes), elle est suivie par celle de Bourgoin-Jallieu puis par la communauté d'emploi de Vienne. Ces communautés sont caractérisées par la présence de communes intermédiaires en lien direct avec le pôle nodal et avec un ensemble de commune placé sous leur influence. Pour caractériser ces lieux intermédiaires, deux approches sont possibles. La première consiste à observer le graphe et la seconde à estimer la position des lieux dans le système d'échanges, par le recours à des mesures locales de centralité.

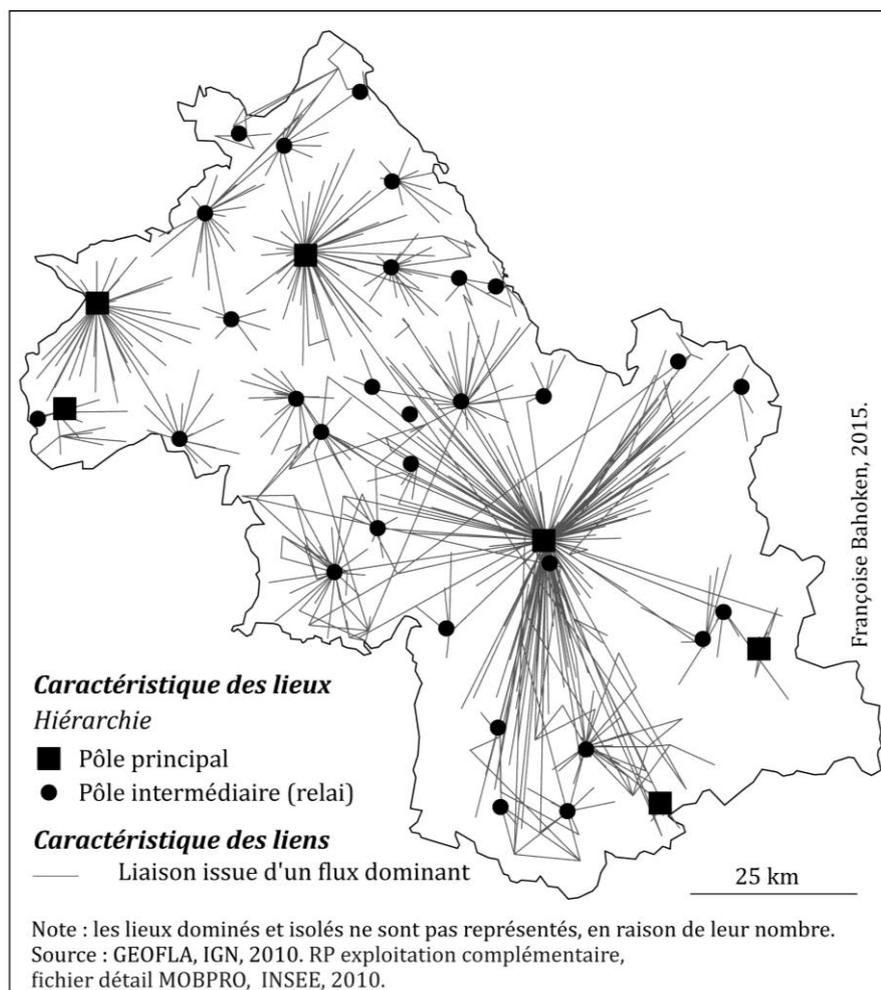
L'approche que nous mettons en œuvre est visuelle et s'appuie sur l'observation de la Figure 9-14.

Effectivement, un bon moyen d'estimer la position de relais d'un lieu consiste à observer le graphe précédent. Ces lieux permettent de faire le lien entre deux catégories de lieux : les dominants et les dominés. Cette position intermédiaire confère aux lieux qu'elles dominent une valeur de distance géodésique de 2 liens qu'il convient de parcourir pour rejoindre Grenoble, tandis que l'ensemble des autres lieux dominés sont à une distance de 1. D'après le graphe de la Figure 9-14, le département de l'Isère comptabilise ainsi 29 lieux relais qui sont à la fois dominés par un pôle, et dominant un ensemble de communes. Ces lieux définissent ainsi un ensemble plus ou moins complexes de sous-régions, en fonction de leur distance géodésique au pôle nodal.

La seconde approche consiste à calculer une mesure estimant la centralité des lieux dans l'ensemble du graphe. Etant donné que plusieurs d'entre elles sont disponibles dans la littérature, nous signalons *a minima* la centralité de degré pondéré qui estime cette position par rapport au nombre de liens, en fonction de leur valeur et de leur orientation, selon qu'ils sont dirigés vers les lieux (pour identifier les pôles dominants, aussi qualifiés « d'autorités » car ils ne sont dominés par aucun lieu), majoritairement sortants des lieux (pôles dominés), ou bien ont une position de relais.

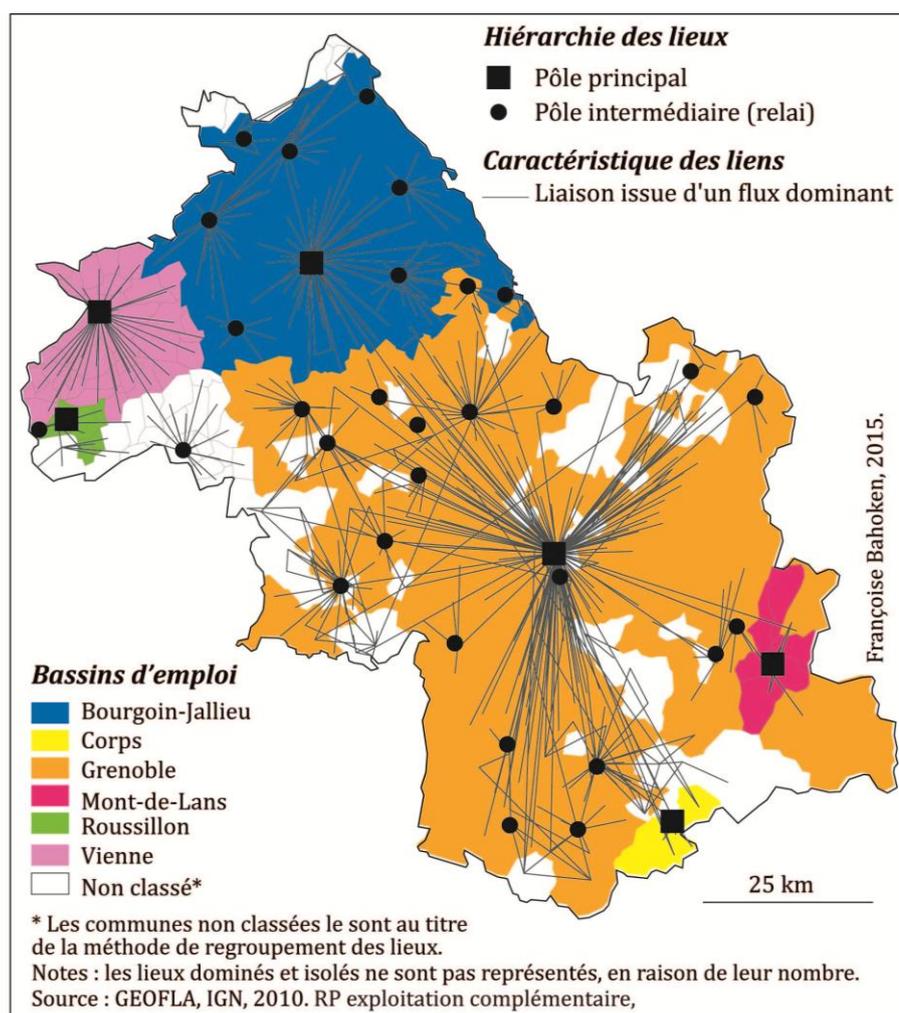
La Figure 9-15 présente une spatialisation du graphe de la Figure 9-14, correspondant à l'Etape 3 de la procédure.

Figure 9-15. Cartographie des pôles nodaux et des sous-bassins d'emploi du système de flux (Fij) dominants (Fid) de l'Isère



Sur le plan analytique : on distingue bien les pôles principaux (les carrés noirs) qui polarisent les liaisons dominantes donnant l'apparence d'une représentation en oursins. Les liens correspondent aux relations issues de l'analyse *de flux dominants*, ils sont dirigés vers les lieux dominants ainsi que vers les relais qui conduisent aux étapes suivantes de formation de régions ou de sous-régions nodales.

La Figure 9-16 présente une analyse visuelle de régionalisation de l'Isère, correspondant à l'Etape 6 de la procédure. Elle met en évidence des bassins d'emplois polarisés par les six pôles nodaux identifiés précédemment. On observe que ces bassins définissent des aires mitées par la présence de communes qui sont non classées au titre de la méthode utilisée pour leur création (Fruchterman-Reingold) : sur les franges du département (au nord, en bordure du département du Rhône ou à l'ouest, à la frontière de la Drôme) mais également au cœur du bassin de Grenoble.

Figure 9-16. Régions et sous-régions du système de *flux dominants* (Fid) de l'Isère (Fij)

La méthode apparaît ainsi adaptée au tracé de communautés sous la forme de graphes, elle ne semble pas l'être pour celle de « communautés spatialisées » qui correspondent à des régions. Sans doute devons-nous voir dans ce résultat l'une de ses limites majeures : la non prise en compte de la géographie des lieux, de leur topographie rendant, de fait, la méthode inadaptée à la formation de *régions géographiques*. Les régions ainsi générées, dont la continuité spatiale n'est pas assurée, correspondent à des « zones » et non à de pures « régions » (à des zones géographiques formées d'un ensemble d'unités spatiales), au même titre que la carte de flux ne contribue pas à inscrire des flux dans un espace géographique, à l'inverse de la carte de mouvements (voir chapitre 6). Ceci étant, le découpage territorial résultant de l'analyse n'est pas erroné dans la mesure où il correspond à peu près à la réalité du découpage territorial de l'isère en zones d'emplois.

Pour valider cette analyse, nous proposons de la comparer avec les Zones d'emplois (ZE) de l'Isère définies par l'INSEE, dans la sous-section suivante.

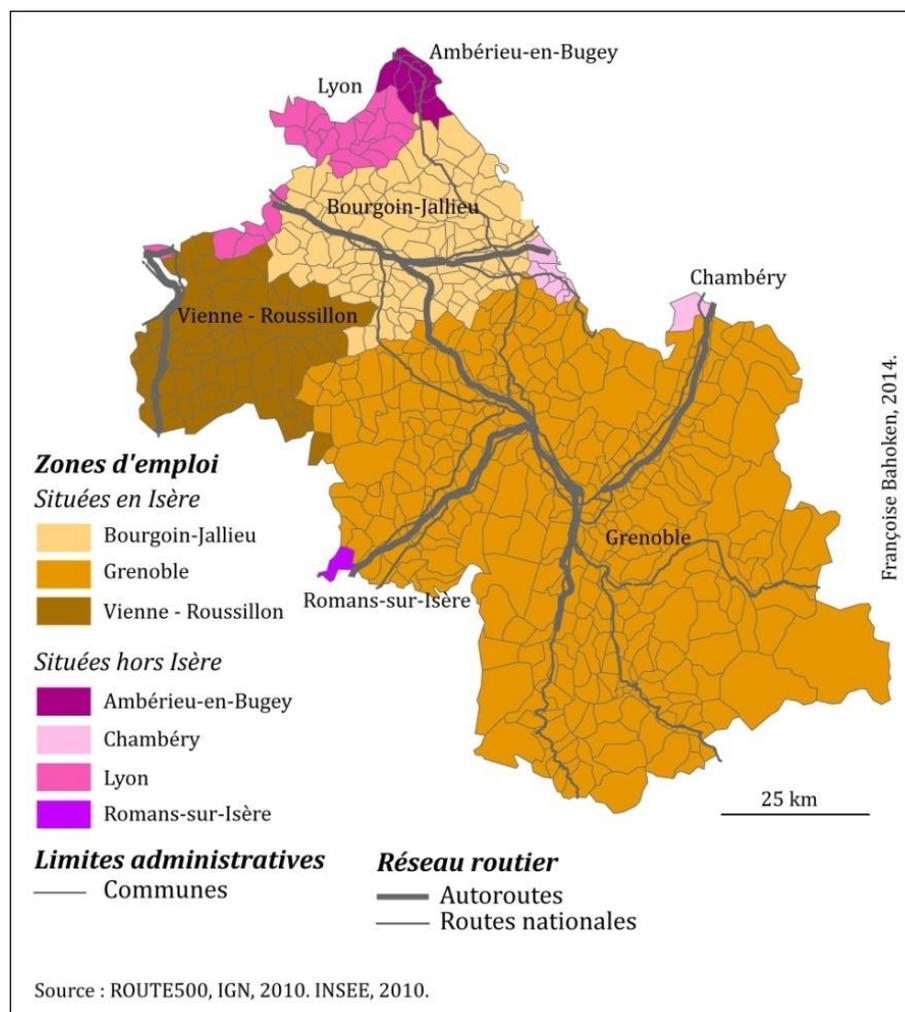
9.2.5.2 Régionalisation en flux dominants versus en Zone d'emplois (INSEE)

Le ZE de l'INSEE correspond à une partition territoriale formées de zones à l'intérieur desquelles les actifs de 15 ans et plus qui y résident y travaillent également. D'après l'INSEE, « *Les zones d'emploi constituent les espaces les plus pertinents pour étudier les problématiques relatives à l'emploi des personnes et à leur lieu de résidence, comme le marché du travail ou les déplacements domicile-*

travail. Elles sont le reflet de la structuration du territoire par la localisation des emplois et des actifs qui les occupent » (INSEE, 2011 : 1).

La Figure 9-17 présente le découpage en ZE de la Rhône-Alpes. Celui est formé de trois ZE principales que sont Bourgoin-Jallieu (n° 8 209), Grenoble (n°8 210) et Vienne-Roussillon (n° 8 211) qui sont partie prenante d'un zonage en 24 zones défini à l'échelle de la région Rhône-Alpes. Elles sont représentées dans les tons bruns sur cette figure.

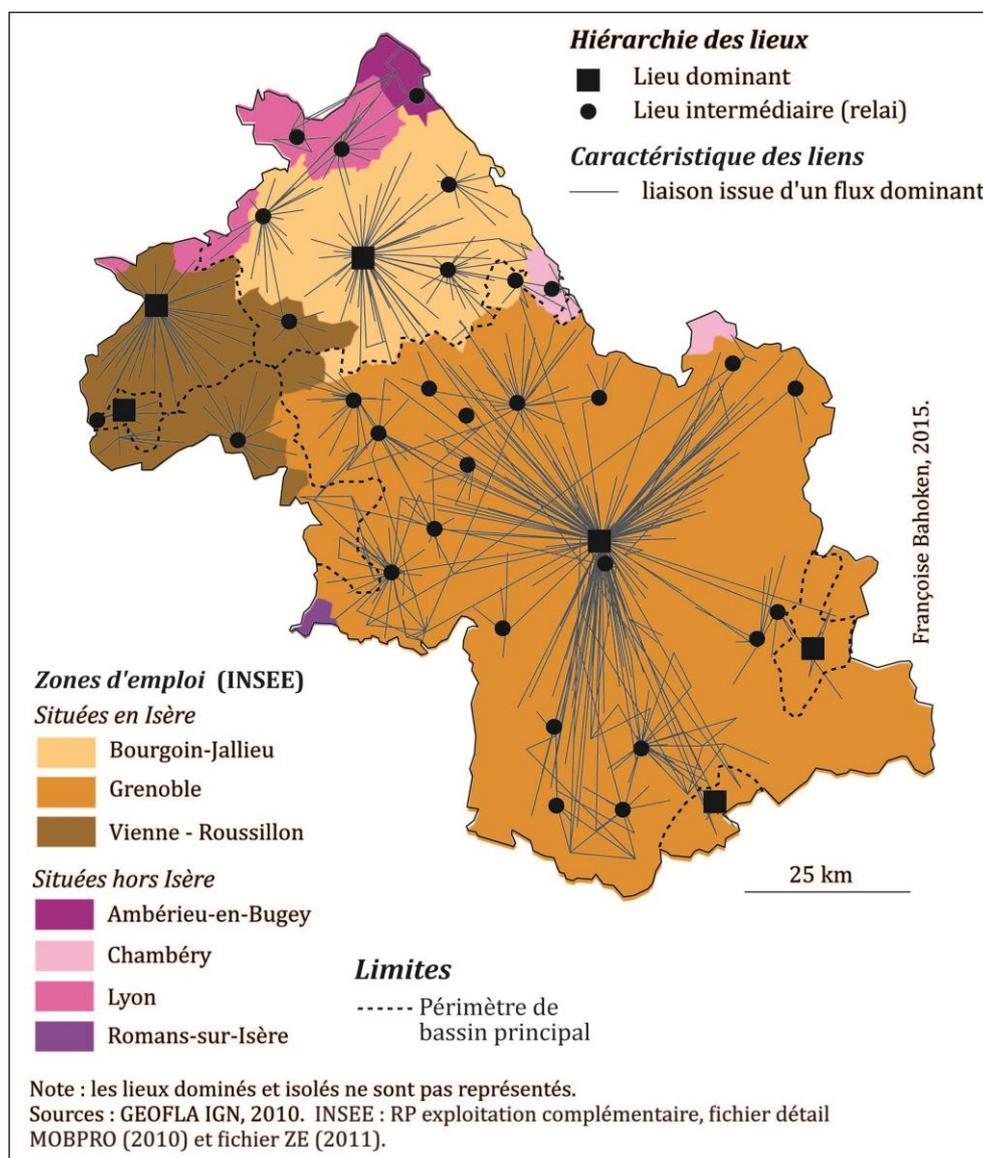
Figure 9-17. Le découpage de l'Isère en Zone d'Emplois, selon l'INSEE (2011).



La partie iséroise du découpage territorial en ZE de la région Rhône-Alpes présente certaines particularités qu'il est important de préciser. Tout d'abord, ce zonage n'est pas totalement inclus dans les limites départementales. Une partie du nord du département est concernée par des ZE qui appartiennent au département voisin du Rhône, c'est le cas des communes situées aux confins des bassins économiques de Bourgoin-Jallieu et de Vienne-Roussillon qui concernent la partie sud-est de la ZE de Lyon et la partie est de la ZE de Saint-Etienne. La raison est liée d'une part, à l'extension spatiale du phénomène de métropolisation de l'agglomération lyonnaise et d'autre part, au fait que les mobilités professionnelles ne sont pas sensibles aux positions des limites administratives communales ou départementales. D'après l'INSEE, « 22 000 actifs résidant dans la ZE de Bourgoin-Jallieu et 19 000 résidant dans celle de Vienne-Roussillon (respectivement 25 % et 22 % des actifs résidant dans ces zones) travaillent dans la ZE de Lyon » (INSEE, 2011 : 2). Ceci est lié au fait que ces ZE sont issues du redécoupage d'une grande ZE antérieure centrée sur Lyon, qui attire encore aujourd'hui 20% des actifs qui résident dans des zones voisines. L'Isère comptant trois ZE principales, la suppression

de celles qui sont situées hors du département n'est pas utile ici, car l'analyse vise à montrer la différence des périmètres de bassins d'emplois obtenus à l'issue d'une analyse des flux dominants par rapport aux ZE de l'Insee. Cette comparaison visuelle est présentée sur la Figure 9-18 qui superpose ces deux informations.

Figure 9-18. Comparaison des principaux bassins régionaux issus des flux dominants (Fid) avec les Zones d'emplois de l'INSEE, pour l'Isère



La superposition sur le même fond de carte, des périmètres de bassins d'emplois obtenus à l'issue de l'analyse des *flux dominants* et des ZE de l'INSEE met en évidence un découpage globalement similaire. On peut toutefois noter quelques différences sensibles (voir Figure 9-16) : en premier lieu, à l'interface de la ZE de Grenoble dans sa partie ouest avec celle de Vienne-Roussillon, certaines communes sont soit non classées soit non intégrées à la ZE de Vienne-Roussillon, de même qu'à l'interface des ZE de Grenoble et de Bourgoin-Jallieu, a distribution des communes dans chacun des bassins et des régions n'est pas similaire. On observe également que l'absence de classement des communes du nord du département (précédemment mentionné pour la Figure 9-16) correspond à

celles qui appartiennent aux ZE situées hors Isère, dans le Rhône ou dans l'Ain. Enfin, les ZE de l'INSEE sont continues dans l'espace et non mitées, à l'inverse du découpage en bassins d'emplois obtenus ; la ZE de Grenoble concerne aussi les trois-quarts sud du département (les bassins locaux de Mont-de-Lans et de Corps ne sont pas considérés comme tels, à l'échelle du département). Toutes ces nuances suggèrent, comme nous l'avons déjà mentionné dans les paragraphes précédents, l'hypothèse **d'une inadéquation de cette méthode de regroupement fondée sur la détection de communautés dans un graphe** spatialisé. La raison tiendrait notamment à l'absence de prise en compte du contexte géographique. Ceci étant, cette affirmation est à manier avec précaution dans la mesure où il est nécessaire de poursuivre l'investigation par l'examen attentif des deux méthodes de regroupement ou de régionalisation utilisées (choix des critères de regroupement, impératifs de continuité spatiale, de contiguïté des zones, ...).

Sur le plan cartographique, la réintroduction de la couleur pour symboliser les différents bassins et zones d'emplois sous la forme d'aplats (plutôt que de points) renforce le message. La commune de Grenoble et sa région sont ainsi bien perçues comme présentant l'extension spatiale la plus importante du département, ce qui correspond au résultat d'une analyse s'inscrivant dans une *logique de lieux* (et les liens) plutôt que de flux. Cependant, la différence de taille du bassin d'emploi de Grenoble, par rapport aux deux autres, est figurée par un aplats de couleur dans le cadre d'un *raisonnement qualitatif*. Cette méthode pourrait générer des biais de lecture émanant d'effets de taille des communes : en termes de superficies mais aussi d'attractivité, les trois ZE iséroises n'étant d'ailleurs pas directement comparables pour ce qui est de leurs capacités attractives (versus la quantité d'emploi proposée). C'est pourquoi il apparaît qu'il faille davantage discuter de la manière de représenter l'attractivité différenciée de ces communes, pour améliorer cette cartographie de la régionalisation en bassins issues de l'analyse des *flux dominants* ; aussi d'envisager les conditions d'une représentation plus évocatrice de la hiérarchie des lieux.

Avant d'aller plus loin, nous proposons de valider la reconstruction cartographique de la méthode Nyusten-Dacey sur une matrice de flux mondiaux.

9.3 Validation empirique : l'exemple de flux commerciaux internationaux (OMC, 2007)

Cette section a pour objectif de valider la reconstruction cartographique de la méthode Nyusten-Dacey d'analyse des *flux dominants*, ainsi que le passage d'une *logique de lieux discrète* à une *logique de lieux continue* – pour cartographier leur attractivité –, en l'appliquant sur une nouvelle matrice de flux internationaux commerciaux, celle de l'OMC (2007).

L'application de la méthode Nyusten-Dacey sur une matrice de flux (Fij) constitue, en réalité, un cas particulier, si l'on considère les différentes possibilités de décomposition de la matrice : (Lij), (Fij), (Fji), (Fij⁺) et (Fij⁻), comme présenté dans le chapitre 3. En effet, les analyses présentées ci-dessus ne focalisent l'attention que sur un seul aspect de la matrice : sur les lieux de destination (j) de (Fij) via leurs relations (Lij) et sur la notion d'attractivité des zones (Dj). Pour une matrice de navettes telle que celle de l'Isère, cela signifie que l'analyse n'est réalisée qu'en termes d'emplois des communes de destination. En effet, la carte de la Figure 9-15 (page 370) met en évidence des zones d'emplois formées autour de lieux dominants ou intermédiaires, jouant le rôle de relais : elle ne décrit ni des zones de résidence ni de peuplement des travailleurs, elle ne met pas en évidence des bassins résidentiels ou fournisseurs d'emplois à proprement parler, car elle est focalisée sur le lieu de destination correspondant, en l'occurrence à la commune du lieu de travail.

La cartographie des *flux dominants* est, de ce fait, susceptible de varier en fonction du type de matrice, selon qu'elle est asymétrique ou non. Cela a pour conséquences, pour les matrices asymétriques, que

l'attractivité peut être observée du point de vue des lieux destination (Dj) comme dans l'approche Nyusten-Dacey, ou d'origine (Oi) en fonction de la thématique du flux. Ce point est d'autant plus important que le graphe sous-jacent de ces matrices asymétriques présente des lieux d'origine qui diffèrent de ceux de destination sur le plan thématique, comme, par exemple, dans le cas de la matrice de l'Isère (commune de résidence * commune d'emploi). De même que le graphe des matrices de flux commerciaux internationaux d'importation (Fji) et d'exportation (Fij) sera différent et la cartographie des flux dominants qui en découle variée.

Tenir compte des différents types de matrices existants et des spécificités thématiques de certaines d'entre elles nous conduit à sortir du cas général de (Fij), pour évoquer les conditions de l'application de la méthode Nyusten-Dacey aux quatre familles de matrices disponibles (chapitre 3). La section précédente ayant présenté le cas de (Fij) appliqué à l'Isère, nous allons maintenant le valider, en exploitant une matrice de flux internationaux commerciaux.

Nous présentons d'abord les données de flux internationaux (OMC, 2007) qui serviront de validation empirique, puis le cas usuel d'application de la méthode sur la composante (Fij) de ces flux commerciaux, avant d'envisager ses différentes possibilités d'application sur (Fji) et (Fij⁺).

9.3.1 Cartographie des flux commerciaux dominants (Fij), le cas officiel

La base de données des flux commerciaux, générée par l'OMC, décrit les flux bilatéraux de marchandises entre les pays membres de l'organisation en 2007. Le choix d'utiliser cette base de données, par rapport à celle de CHELEM qui porte sur des données similaires, tient à des raisons thématiques et méthodologiques combinées. Les données n'ayant pas fait l'objet d'une agrégation spatiale et/ou temporelle (voir section 5.4), la couverture spatiale est de ce fait plus importante. Les valeurs de flux sont effectivement disponibles entre les pays de l'organisation, sans que certains d'entre eux soient regroupés sur une base géographique ou non (zones de libre-échange, unions douanières, ...). Sur le plan géographique, le fond de carte associé à ces flux internationaux de marchandises est par conséquent celui du découpage mondial ordinaire en états nations reconnus.

Ce choix de flux de l'OMC se justifie surtout par le fait que *l'effet de taille* des lieux serait faussé avec CHELEM, car l'analyse mélangerait des états et des agrégats d'états, ces derniers ne disposant pas d'une autonomie économique et d'une légitimité politique.

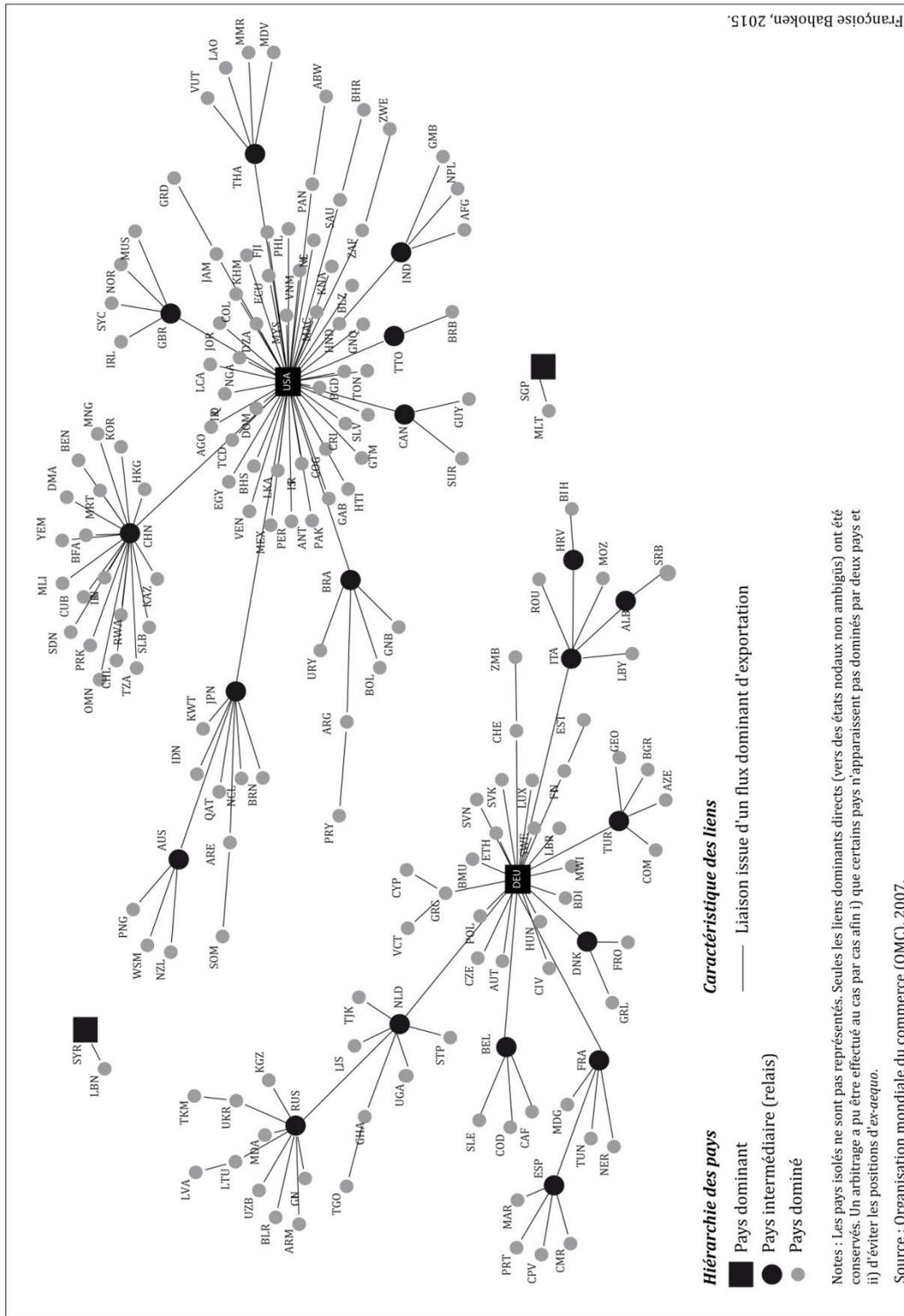
Les données de l'OMC sont issues de la comptabilité douanière ; elles sont exprimées en US \$ et portent sur l'année 2007. Les exportations sont exprimées en base franco à bord (f.a.b.) et les importations sont exprimées en coût, assurance et fret (c.a.f). Ces données ont été préparées par G. Van Hamme, dans le cadre du projet *EuroBroadMap* déjà mentionné.

La première étape de la procédure consiste à cartographier les *flux majeurs*, en mobilisant la *Taille* des liens, mais elle ne sera pas présentée pour des raisons évoquées précédemment : la cartographie des *flux majeurs* correspond à une variante des cas précédents mobilisant un critère de sélection des valeurs et la *largeur* du figuré. Nous allons plutôt explorer les variations de la cartographie de l'attractivité des lieux, au regard du commerce mondial, en représentant les liaisons internationales issues d'une analyse des *flux dominants*.

La première étape principale consiste donc à cartographier les *flux dominants* (fid) de (Fij) selon la méthode Nyusten-Dacey (Etape 1 à 3 décrites précédemment) : elle conduit à explorer le graphe correspondant puis à le spatialiser en le projetant sur un fond de carte polaire. La seconde étape principale est relative à la régionalisation en systèmes et sous-systèmes du commerce mondial.

La figure classique des flux commerciaux dominants est celle du graphe nodal correspondant à l'étape 3 de la procédure (voir supra page 362), il est illustré sur la Figure 9-19.

Figure 9-19. Graphe des liaisons commerciales nodales issues de l'analyse des *flux dominants* (Fij) (méthode Nyusten-Dacey, 1961)



Ce graphe des relations dominantes de l'OMC (2007) est obtenu suite à un repositionnement effectué avec la méthode²¹⁴ de T. Fruchterman et E. Reingold (1991). Réalisée dans un style en noir et blanc, elle mobilise une implantation ponctuelle des pays, associée à des variations de *Forme* pour représenter leur hiérarchie en trois classes de lieu (dominant, relais, dominé) ; les liens représentés sont ceux qui résultent de l'analyse des flux dominants.

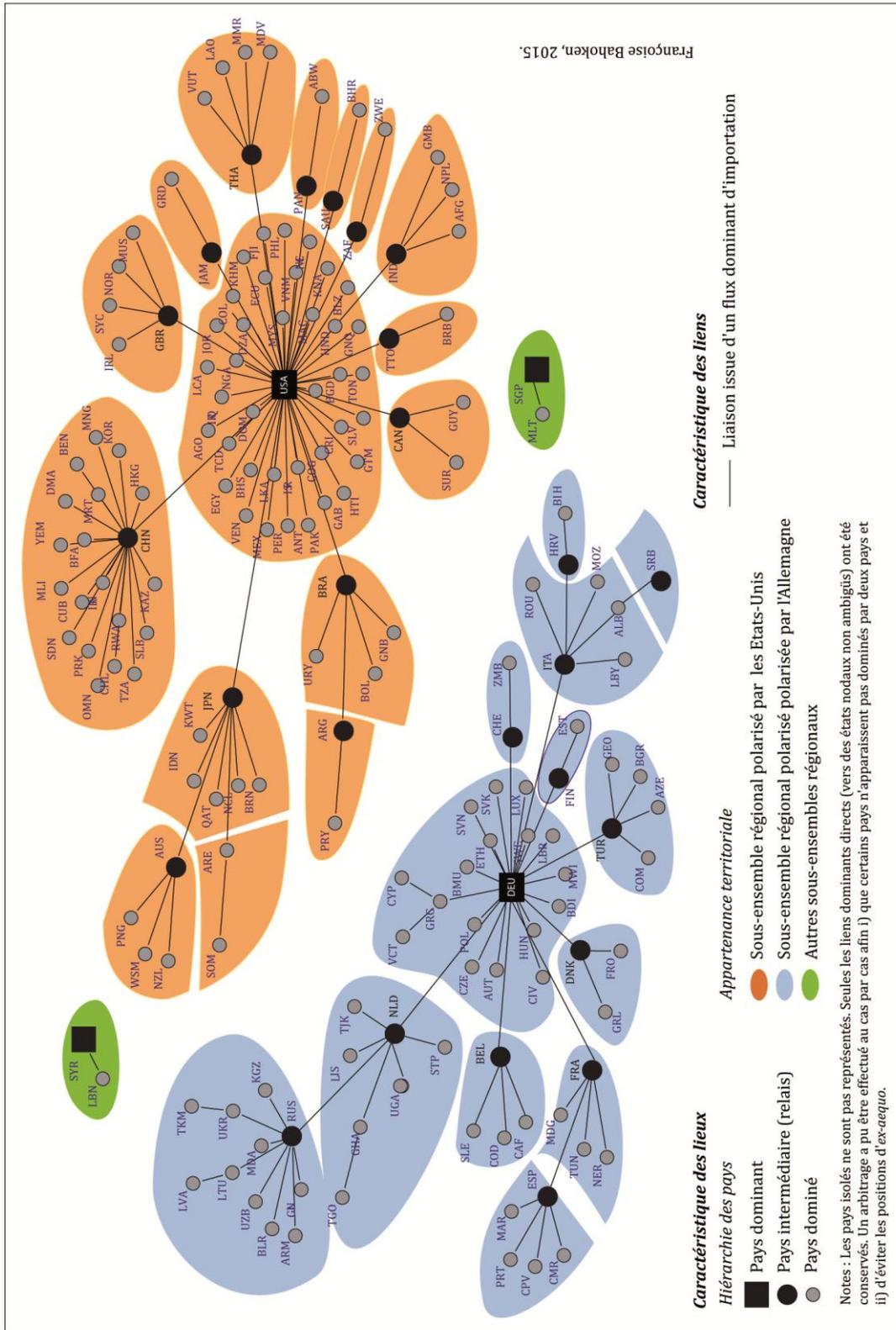
Il est important de noter que les liaisons représentées, issues de flux dominants de (Fij), mettent en évidence un flux d'importation et, plus généralement, les deux grandes régions réceptrices de flux commerciaux exprimés en valeur. La raison tient au fait que le raisonnement mené pour caractériser la hiérarchie des lieux est réalisé sur (Dj), la somme marginale en colonne dont on rappelle qu'elle traduit la capacité de réception des lieux, donc d'importation. De ce point de vue, ce graphe décrit des groupes de pays qui reçoivent des valeurs de leurs vendeurs (ceux qui exportent).

A un second niveau de lecture, on observe que ces deux régions principales sont décomposables en sous-régions, comme illustré sur la Figure 9-20. Le monde apparaît ainsi scindé en deux principales régions importatrices polarisées par l'Allemagne (DEU) et par les Etats-Unis (USA), auxquelles il est possible d'ajouter deux petits groupes d'états correspondant aux liens entre la Syrie et le Liban (SYR-LBN) et entre Malte et Singapour (MLT-SGP), hormis les pays isolés – au sens de l'algorithme de positionnement utilisé – qui ne sont pas représentés.

Toujours sur la Figure 9-20, ces deux régions correspondent à un découpage de niveau 1. Elles apparaissent structurées autour de pays intermédiaires qui présentent une position de relais, autorisant la formation de sous-régions correspondant à un découpage de niveau 2. Ces pays relais sont, en effet, des fournisseurs (ou exportateurs) du pôle de niveau 1 qui leur importe des produits, tout en étant eux-mêmes importateurs de valeurs monétaires auprès d'un chapelet de pays du point de vue desquels ils sont exportateurs. Dans certains cas, ce chapelet de pays entraîne la formations de niveaux 3 voire 4 de découpage, comme dans le cas de la macro-région polarisée par l'Allemagne.

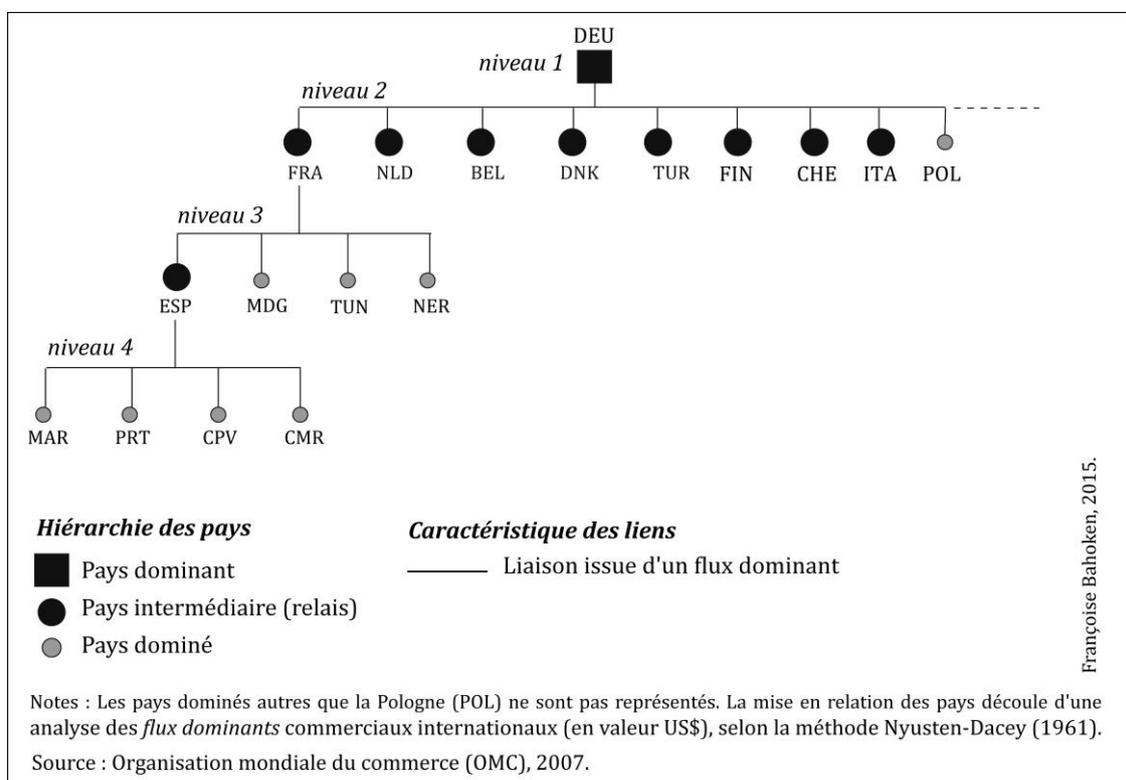
²¹⁴ Nous rappelons que cet algorithme adapté aux grands graphes est fondé sur des caractéristiques nodales d'attractivité ou de répulsion des nœuds ainsi que sur une fonction de type masse-ressort, analogue à la gravitation, appliquée aux liens. Ce dernier entraîne la formation de groupes et sous-groupes ou communautés de nœuds.

Figure 9-20. Graphe des régions et sous-régions nodales issues de l'analyse des flux dominants du commerce mondial d'importation (Fij) de l'OMC (2007)



Pour analyser cette hiérarchie finement, il est possible de se référer au graphe précédent ou de le représenter sous la forme d'un arbre dont la lecture plus aisée peut apparaître complémentaire. La Figure 9-21 présente à titre d'exemple la hiérarchie de la région (DEU).

Figure 9-21. Arborecence principale des pays formant la région polarisée par l'Allemagne



Sur les 46 pays formant la région (DEU), huit d'entre eux forment des sous-régions de niveau 2 : la France (FRA), les Pays-Bas (NLD), la Belgique (BEL), le Danemark (DNK), la Turquie (TUR), la Finlande (FIN), la Suisse (CHE), l'Italie (ITA) ; la Pologne (POL) est un état dominé placée au niveau 2. La sous-région (FRA) peut être subdivisée en deux niveaux formant huit pays depuis desquels elle importe directement : l'Espagne (ESP), Madagascar (MDG), la Tunisie (TUN) et le Niger (NER) ou indirectement, par le biais du relais espagnol : le Maroc (MAR), le Portugal (PRT), le Cap-Vert (CPV) et le Cameroun (CMR). La macro-région américaine (non représentée ici) est quant à elle formée des régions polarisées par la Grande-Bretagne (GBR), la Chine (CHN), le Japon (JPN) et l'Australie, le Brésil (BRA), le Canada (CAN), Trinité-et-Tobago (TT0), l'Inde (IND), le Panama (PAN), l'Arabie-Saoudite (SAU), l'Afrique du sud (ZAF) et la Thaïlande (THA).

Sur le plan thématique, l'analyse de ces graphes ne doit pas perdre de vue qu'ils représentent une sélection de l'information, qui s'appuie essentiellement sur capacités attractives des lieux (le critère [9.2]).

La question du choix du premier critère de sélection global [9.1], conduisant au *flux majeur*, est également importante (voir Chapitre 8), car elle entraîne le creusement de la matrice. Si l'on peut considérer exact que la France domine la Tunisie, en correspondant à son premier marché, il n'est pas impossible que des flux de valeurs similaires ou supérieures à celui la soient dirigés vers la France sans pourtant autant être représentés. La raison tenant à ce qu'ils ne sont pas de rang 1. Il faut donc garder à l'esprit que la méthode correspond bien à la représentation d'une sélection très forte des seules relations commerciales qui sont issues d'une analyse des *flux dominants* (méthode Nyusten-Dacey).

Du point de vue de l'interprétation du résultat, la hiérarchie des lieux, présentée sur la Figure 9-19, fait apparaître deux régions organisées autour de l'Allemagne et des Etats-Unis, ce qui pose question quant à l'absence de la Chine. Plusieurs hypothèses qui tiennent à la sémantique de ces flux peuvent être avancées.

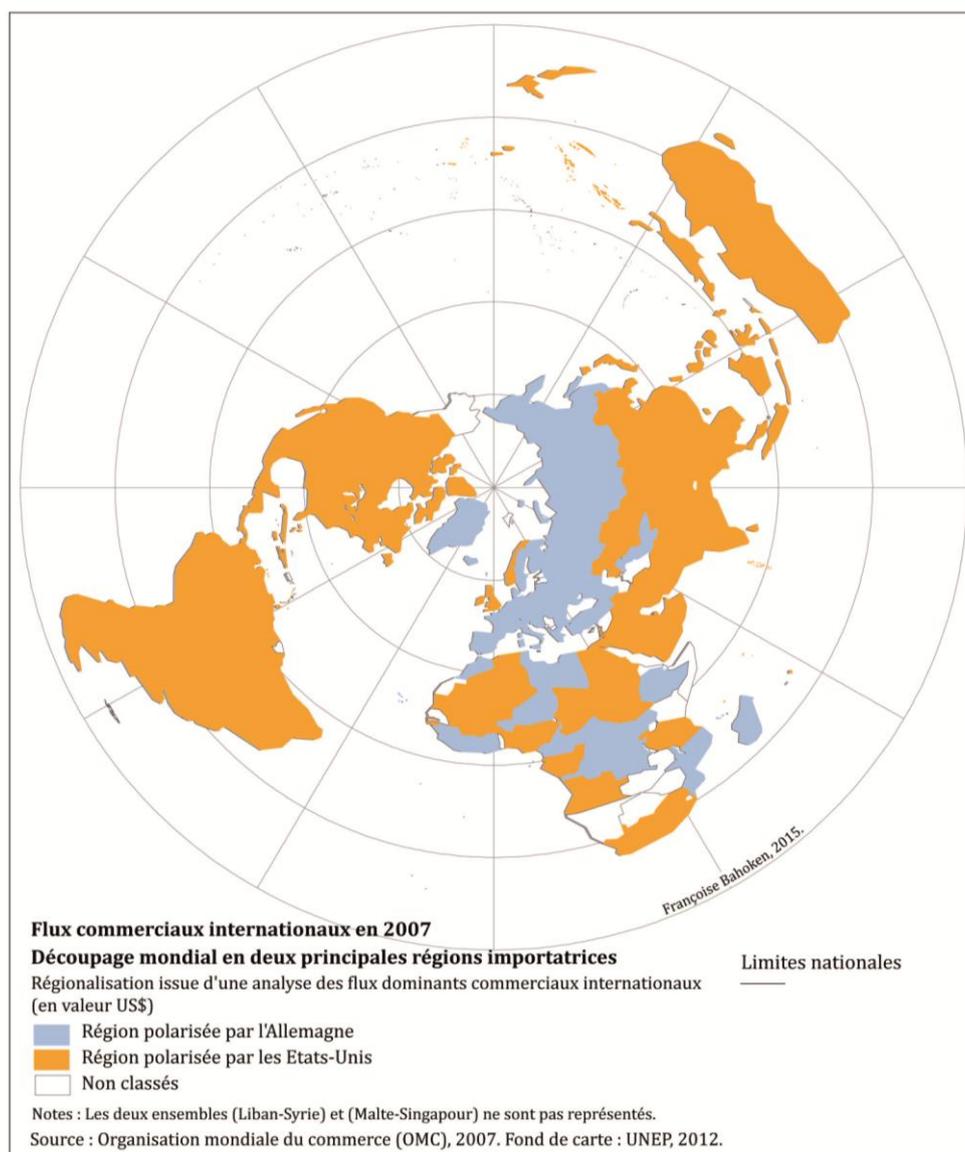
La Chine, bien connue pour être « l'usine du monde », serait en fait un pays qui importe de l'étranger des produits à faible valeur monétaire d'où son absence en tant que grande région importatrice de valeurs. Ces importations seraient essentiellement des matières premières qu'elle transformerait ensuite en biens à forte valeur qui en ressortiraient sous la forme d'exportations. Prenons un exemple : la Chine importerait de la ferraille de faible valeur et la transformerai en véhicule de plus forte valeur monétaire qui seraient ensuite exportées de manière préférentielle vers l'Europe. La Chine étant un grand pays, certaines matières premières utiles à son activité productive seraient issues de transfert internes : ses usines seraient en fait approvisionnées à partir de matières premières qui seraient présentes dans l'une de ses régions d'extraction de minerais de fer ou de charbon, ... Ces deux hypothèses expliqueraient l'absence de la Chine du groupe des pays fortement importateurs de biens exprimés en valeur, puisque ses importations seraient soit peu onéreuses (la ferraille), soit proviendraient de flux internes n'apparaissant pas dans les échanges internationaux. En revanche, ces flux internes ressortiraient ensuite sous la forme de flux externes d'exportations onéreux en valeur (des véhicules) ce qui suggère l'hypothèse de la présence d'une région Chine dans la régionalisation correspondant aux flux d'exportations, nous y reviendrons dans la section 9.3.2. Sont également susceptibles d'intervenir dans l'explication de l'absence de la Chine comme grande région importatrice : l'unité de mesure de ces flux, la valeur et non le tonnage ainsi que, probablement, un effet de taille ; la ferraille étant en effet supposée peser plus lourd que de petits composants informatiques, très légers et très onéreux.

La dernière étape de l'analyse consiste en la cartographie de la régionalisation observée sur le graphe précédent. Compte tenu de l'importance de l'arborescence et du découpage pouvant aller jusqu'à quatre niveaux, la cartographie exhaustive n'est pas très intéressante ; elle l'est d'ailleurs d'autant moins que l'on peut montrer que la méthode de régionalisation mise en œuvre n'est probablement pas la plus efficace²¹⁵. P. Dumolard avait d'ailleurs souligné cette limite de l'approche de Nyusten-Dacey (1961) qui ne procurait pas de méthode de régionalisation en tant que telle.

La Figure 9-22 présente le découpage mondial de niveau 1 en deux grandes régions.

²¹⁵ Cela tient au choix de l'algorithme de placement des lieux utilisé pour réaliser les graphes, dont l'objectif est de visualiser des communautés d'acteurs liés aux relations considérées. Il semble optimiser pour cela davantage le placement des nœuds que la régionalisation à proprement parler qui nécessite la prise en compte des positions dans l'espace.

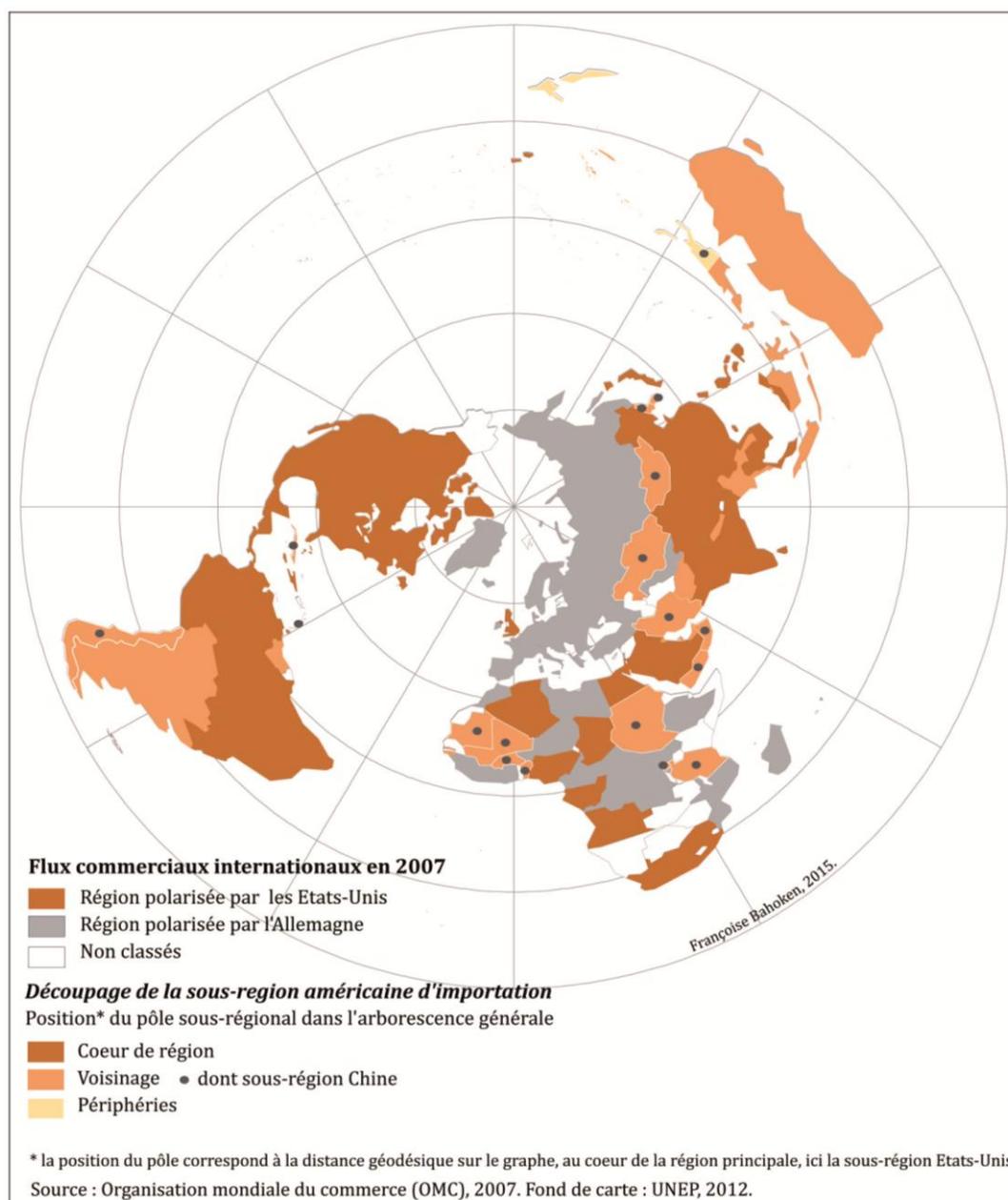
Figure 9-22. Cartographie des deux principales régions mondiales issues des flux dominants (Fij) commerciaux internationaux



Là où la carte présente un intérêt par rapport au graphe (voir Figure 9-19 / Figure 9-22), c'est dans la visualisation de la structure régionale qui découle de l'analyse du graphe, plutôt que dans la visualisation exhaustive des liens. Raisonner ainsi permet de souligner la complémentarité du graphe et de la carte au titre de l'aide à l'interprétation de la structure hiérarchique des flux, lorsqu'un nombre conséquent de lieux est représenté. Si le graphe permet de visualiser instantanément les différents niveaux hiérarchiques de l'arborescence et la position des pays (dominante, de relais ou dominée) lorsque les pays sont placés de sorte à mettre en évidence des communautés d'entre eux pays, là où la carte (voir ci-dessus) apparaît souligner l'inscription spatiale de l'aire d'influence commerciale de principales régions. C'est pourquoi il nous est apparu plus judicieux de focaliser l'analyse cartographique sur le découpage mondial de niveau 1, puis d'examiner la formation de l'une ou de l'autre région aux niveaux suivants.

La Figure 9-23 examine ainsi la formation de la région polarisée par les USA (le choix de la région américaine tenant à la présence de la Chine au niveau 2 du découpage).

Figure 9-23. Cartographie du découpage en sous-régions de l'aire d'influence américaine issue des flux dominants (Fij) commerciaux internationaux

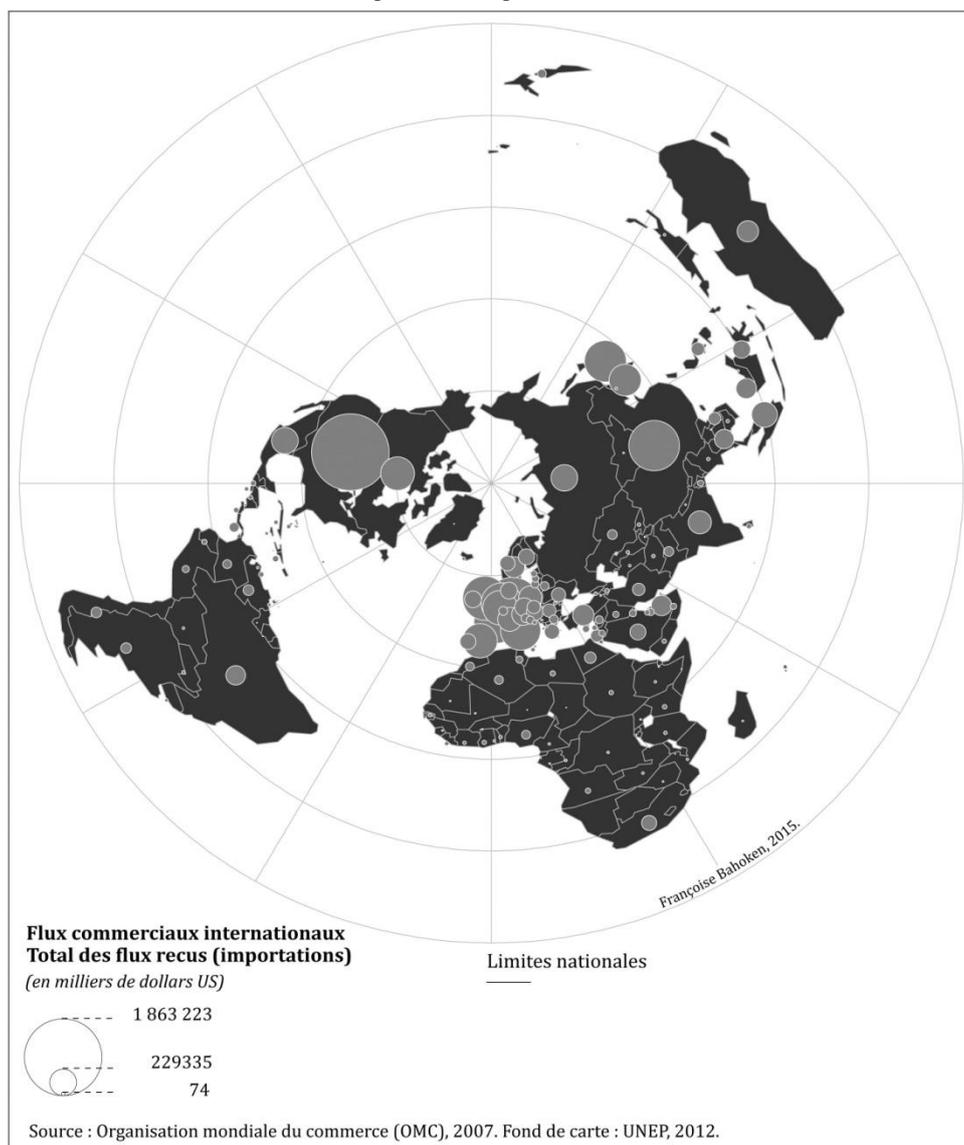


Sur la Figure 9-23, la nuance de teinte des pays est fonction de leur distance géodésique sur le graphe correspondant, en l'occurrence de leur éloignement à la sous-région centrale (les USA). Ainsi, le « cœur de la région » englobe les pays qui sont en liaison directe avec les USA, ceux depuis lesquels les USA importent directement, sans intermédiaire. Le « voisinage » correspond aux pays non directement liés aux USA, ils sont à une distance géodésique de 2 correspondant au niveau 2 de l'arborescence, impliquant un état relais ; les pays situés dans la « périphérie » concernent les niveaux 3 et 4 de l'arborescence. On observe ainsi que la majorité des pays de la région américaine exportent des valeurs soit directement vers les USA, soit en passant par un intermédiaire principal qui se révèle être la Chine ; les autres pays relais apparaissent en outre comme plus spécialisés que la Chine.

Cette carte ne représente pas tant l'attractivité des pays, difficilement lisible à l'aide d'aplats, que l'extension spatiale de leur dépendance aux pays desquels ils importent des biens. La dépendance est

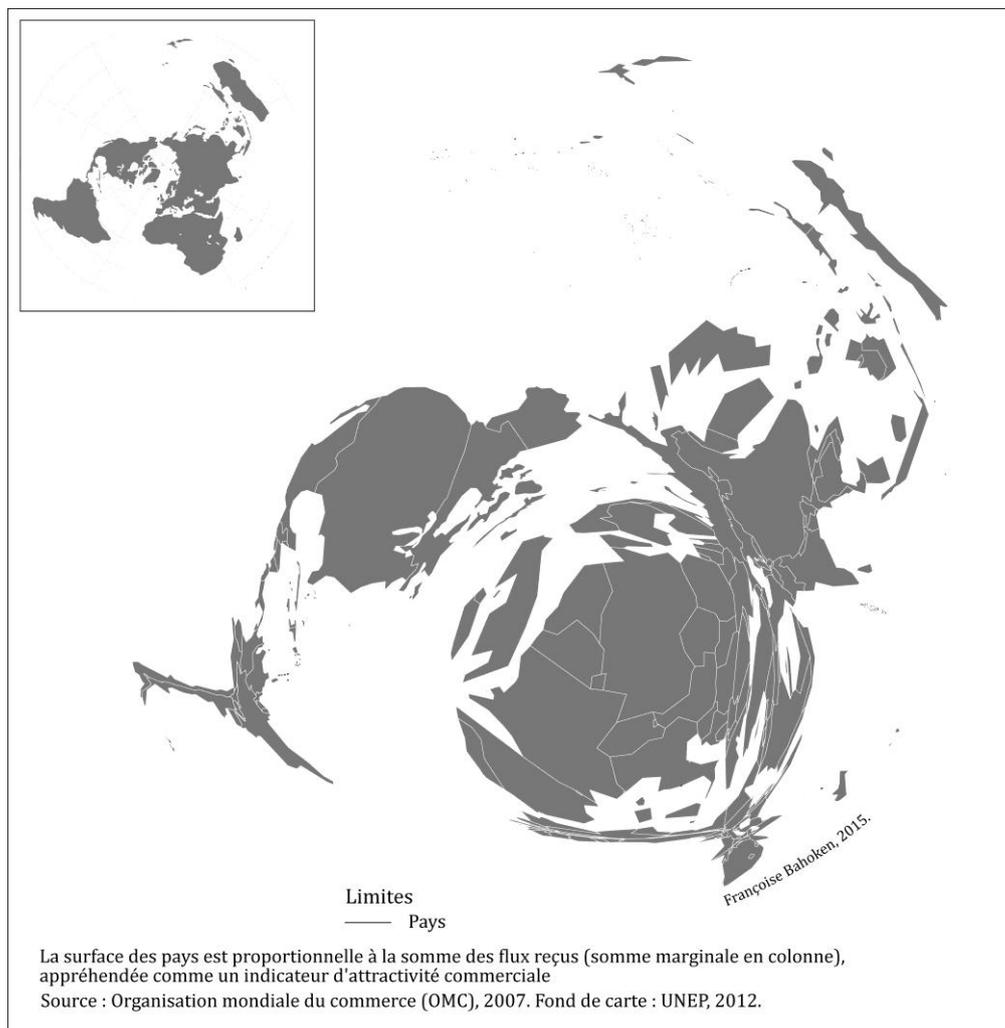
d’ailleurs probablement biaisée par un effet de taille des pays bien connu en cartographie (les pays présentant les plus grandes superficie apparaissent les plus visibles, sans pour autant être les plus importants en valeur). C’est pour remédier à cet éventuel biais pouvant être préjudiciable à l’analyse de la hiérarchie des lieux, par conséquent de leur leur attractivité, que l’on recourt à la *Taille*, en implantation ponctuelle ou surfacique. Deux possibilités sont envisageables, selon que l’implantation est ponctuelle ou surfacique, elles sont respectivement présentées sur la Figure 9-24 et la suivante.

Figure 9-24. Cartographie de l’attractivité des lieux (Dj), par le recours à la *Taille*
Implantation ponctuelle



La cartographie de l’attractivité des lieux sous la forme d’une carte choroplèthe classique souffrant fortement de *l’effet de convergence*, en particulier au cœur de l’UE, on observe qu’il *est a priori* plus intéressant de poursuivre l’investigation sur la représentation de l’attractivité des flux dominants sur le fond anamorphosé, comme illustré par la figure suivante.

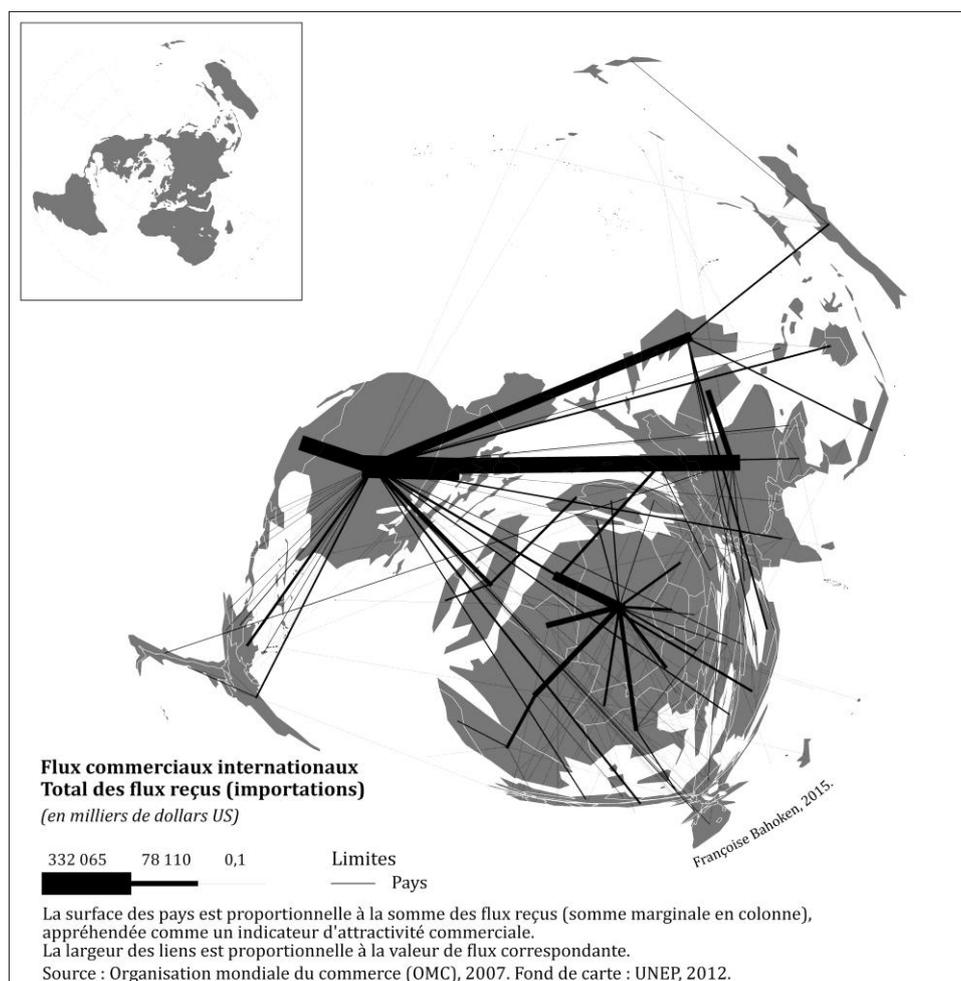
Figure 9-25. Cartographie en anamorphose de la capacité d'importation (Dj) des pays
Implantation surfacique



Cette représentation anamorphosée de l'attractivité (Dj) des pays met clairement en évidence leur capacité de réception, c'est-à-dire leur poids dans le commerce mondial d'importation de biens présentant une valeur financière élevée. Elle met en évidence la position des régions polarisées par les États-Unis et l'Allemagne plus généralement les pays d'Europe occidentale, qui importent beaucoup dans l'ensemble en valeur, mais aussi de la Chine qui n'est pas en reste, en se distinguant nettement des autres pays du sud-est Asiatique.

Il est dès lors intéressant de visualiser la valeur de ces *flux dominants* (et non les seules liaisons, comme le suggère la méthode Nyusten-Dacey), sur ce fond de carte afin de disposer d'une double vision des flux du point de vue des liens (valués) et de la hiérarchie des lieux. Cette solution est présentée sur la Figure 9-26, correspondant d'ailleurs à une proposition récurrente de D. Dorling (1991), déjà mentionnée par la Figure 2-30 (page 99).

Figure 9-26. Cartographie des valeurs de *flux dominants* (Fij) du commerce international, sur un fond de carte anamorphosé



Cette carte représente les *flux* de manière usuelle, selon une *logique de flux* (en mobilisant la *Taille* sur la *largeur* des liens). Sur le plan méthodologique, elle conduit à sortir du cadre strict de la méthode (Nyusten-Dacey) car elle appréhende le second paramètre des auteurs comme un critère de sélection global au titre du Chapitre 8, mais en agissant en outre sur une focale de la matrice : les lieux de destination (les valeurs situées en colonne).

Si cette pratique n'est pas erronée, elle correspond, toujours sur le plan méthodologique, à une variante de l'approche classique au même titre que la cartographie des *flux majeurs* (Fim) de l'Etape 1 de la procédure générale. Cette variante n'est pas dénuée d'intérêts sur le plan thématique : elle présente l'avantage non négligeable de faciliter l'interprétation thématique du résultat obtenu.

En effet, la carte précédente met bien en évidence les relations qui se produisant aux niveaux 1 et 2 de la régionalisation de flux commerciaux mondiaux. Il est aisé de distinguer les pays formant la Triade ainsi que, par exemple, la position de relais de la Chine, celle du Japon dont on observe l'importance des exportations vers les Etats-Unis. Cette forte quantité des importations japonaises concerne certainement les produits technologiques à très forte valeur. On observe également comme un effet de loupe des pays de l'Europe occidentale qui traduisent l'importance de ses relais, parmi lesquels la France, la Belgique, les Pays-Bas mais également la Turquie.

L'analyse des *flux dominants* (Fij) de l'OMC, selon la méthode Nyusten-Dacey (1961), décrit une structure hiérarchique composée de deux principales régions d'importation de produit à forte valeur, à

l'échelle mondiale, au sein desquelles il est possible de considérer la place de la Chine. Elle traduit en effet une forme de dépendance économique de pays fortement importateurs (pays situés aux niveaux 1 et 2 de l'arborescence) vis-à-vis des exportateurs. Cette dépendance interprétée sur le plan thématique est directement liée d'une part, au type de matrice (Fij) et de l'autre, aux critères de sélection retenus.

Ces régions d'importation peuvent ainsi être appréhendées comme dominant les régions productrices et non seulement comme placées sous la domination ou sous l'influence des régions exportatrices. Cette réversibilité de l'interprétation a été bien montrée par A. Bopda (1997) dans sa thèse. L'auteur y explique notamment le fait que certaines régions en dominant d'autres parce qu'elles attirent beaucoup d'acheteurs qui en sont dès lors dépendants, ou qu'elles fournissent beaucoup aux vendeurs. Le fait que l'interprétation de la relation de domination puisse être inversée traduit l'importance de la focale d'analyse des flux dans l'analyse et, plus généralement de l'effet du type de matrice utilisé.

Les analyses présentées ont toutes été réalisées sur la matrice (Fij) dont on a vu qu'elle permettait de visualiser la position des pays du point de vue des importations. Cette focale correspond en pratique au cas officiel d'application de la méthode (Nyusten-Dacey), d'ailleurs toujours utilisé dans la littérature, ce qui pose question. Etant donné les possibilités de décomposition de cette matrice que nous avons examinées dans le chapitre 3, il est en effet possible de considérer ce cas officiel (Fij) mais également d'une part, un cas légal (Fji) et d'autre part, un cas dont on verra qu'il est évident : celui de (Fij⁺). Il convient de noter que ces différentes possibilités ont été mentionnées par les auteurs comme autant de possibilités de sélectionner le « flux le plus important » (« *largest flow* ») émis par chacun des lieux, sous la forme suivante.

« La relation de type 'flux le plus important' peut avoir plusieurs définitions, comme celle du plus large flux sortant [Fij], entrant [Fji] ou total [que l'on suppose correspondre à [Fij⁺]]. Le présent exemple utilise le nombre de flux téléphoniques [Fij] émis par chacune des villes vers toutes les autres de la zone d'étude (Nyusten et Dacey, 1961: 33).

On observe toutefois que les auteurs n'ont pas évoqué le cas de la cartographie des *flux dominants* de (Fij⁻), lequel correspond effectivement à un cas limite. Son utilisation apparaît compliquée, elle est d'autant plus sensible que son analyse nécessite de nombreux arbitrages, car la matrice (Fij⁻) est toujours théoriquement représentée sous une forme simplifiée par l'une des deux cartes d'intérêt correspondant à (Fij⁻)>0 ou à (Fij⁻)<0 dont le choix est fonction de la thématique. Son utilisation génère aussi de nombreuses difficultés d'applications et d'interprétations qui nous conduisent à ne pas nous y intéresser dans l'immédiat.

Aussi allons-nous envisager, dans les deux sections suivantes, les différentes possibilités liées aux matrices de type (Fji) et (Fij⁺), l'hypothèse d'une **variation de l'analyse des flux dominants en fonction du type de matrice étant très forte.**

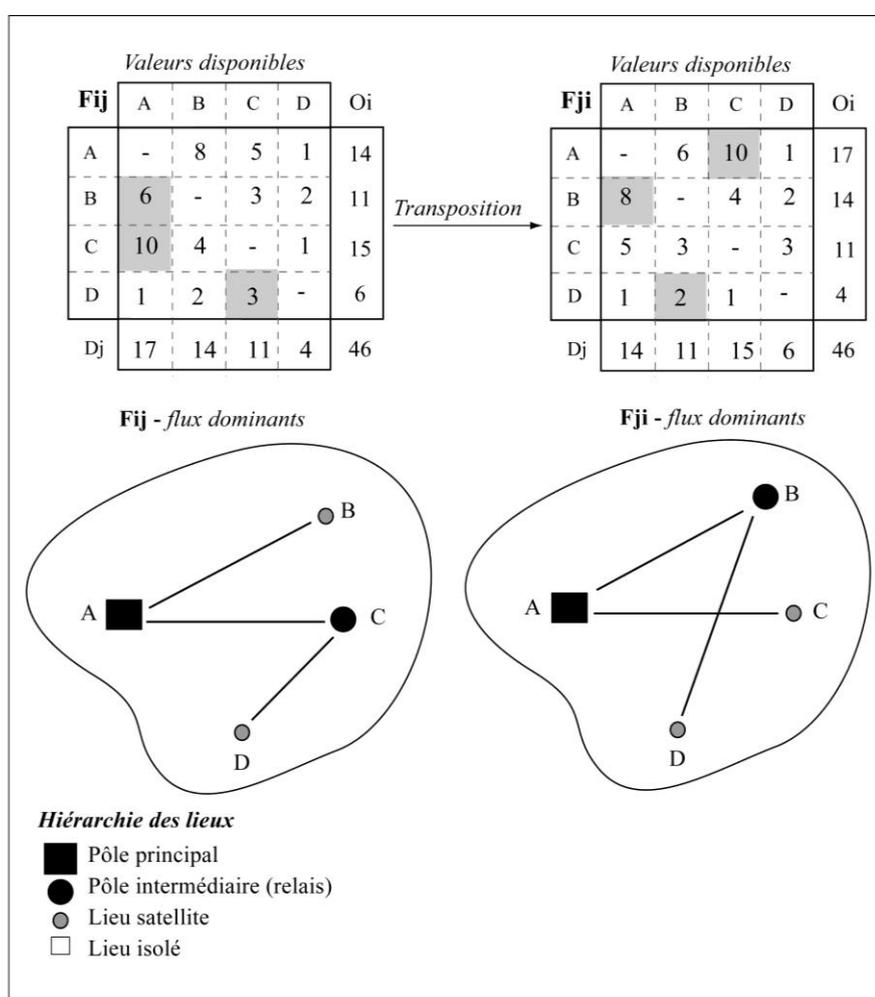
9.3.2 Le cas légal : la cartographie des flux dominants (Fji)

La matrice (Fji), en tant que transposée de (Fij) : (Fji) = ^T(Fij) correspond à ce que l'on pourrait qualifier de cas légal d'application de la méthode Nyusten-Dacey. En effet, elle correspond peu ou prou au cas officiel puisqu'il s'agit de sa transposée. De plus, la matrice demeure asymétrique et décrit une relation également orientée, entre les mêmes lieux que (Fij). La différence entre les deux matrices réside dans le changement de focale lié à la transposition des valeurs, par conséquent à un changement de point de vue thématique et d'orientation des flux. Si l'application pratique de la méthode à (Fji) ne présente pas de différence d'ordres cartographiques et méthodologique avec celle de (Fij), la carte obtenue est en revanche susceptible de varier fortement sur le plan sémantique, en fonction de la thématique du flux. C'est pourquoi l'analyse des *flux dominants* est, en réalité, à manier avec précaution sur les matrices asymétriques, car elles ne sont pas toujours supposées décrire une quantité

de flux entre des lieux qui sont de même nature, par exemple (résidence*résidence). Dans l'exemple des flux commerciaux de l'OMC (2007), le choix de représenter les *flux dominants* à partir de l'exploitation de (Fij) signifie que le regard est porté à l'étape 1 sur les émissions (*flux majeur*), puis à l'étape les liaisons issues des *flux dominants* interprétés en termes de réception (Dj), enfin, aux étapes 4 et 5 sur l'attractivité des lieux importateurs. A l'inverse, le choix d'analyser les *flux dominants* du commerce mondial du point de vue de (Fji) signifie que l'attention est portée, à l'étape 2, sur la capacité exportatrice des pays et à l'étape 5 sur l'aire de marché des pays fortement exportateurs en valeurs financières, sur l'extension spatiale de leurs débouchés économiques.

Le point de vue étant différent entre (Fij) et (Fji) puisque la matrice est asymétrique, l'analyse de la hiérarchie des lieux dans le système des *flux dominants* variera nécessairement localement. La Figure 9-27 illustre cette variation, elle correspond à l'Etape 3 de la procédure.

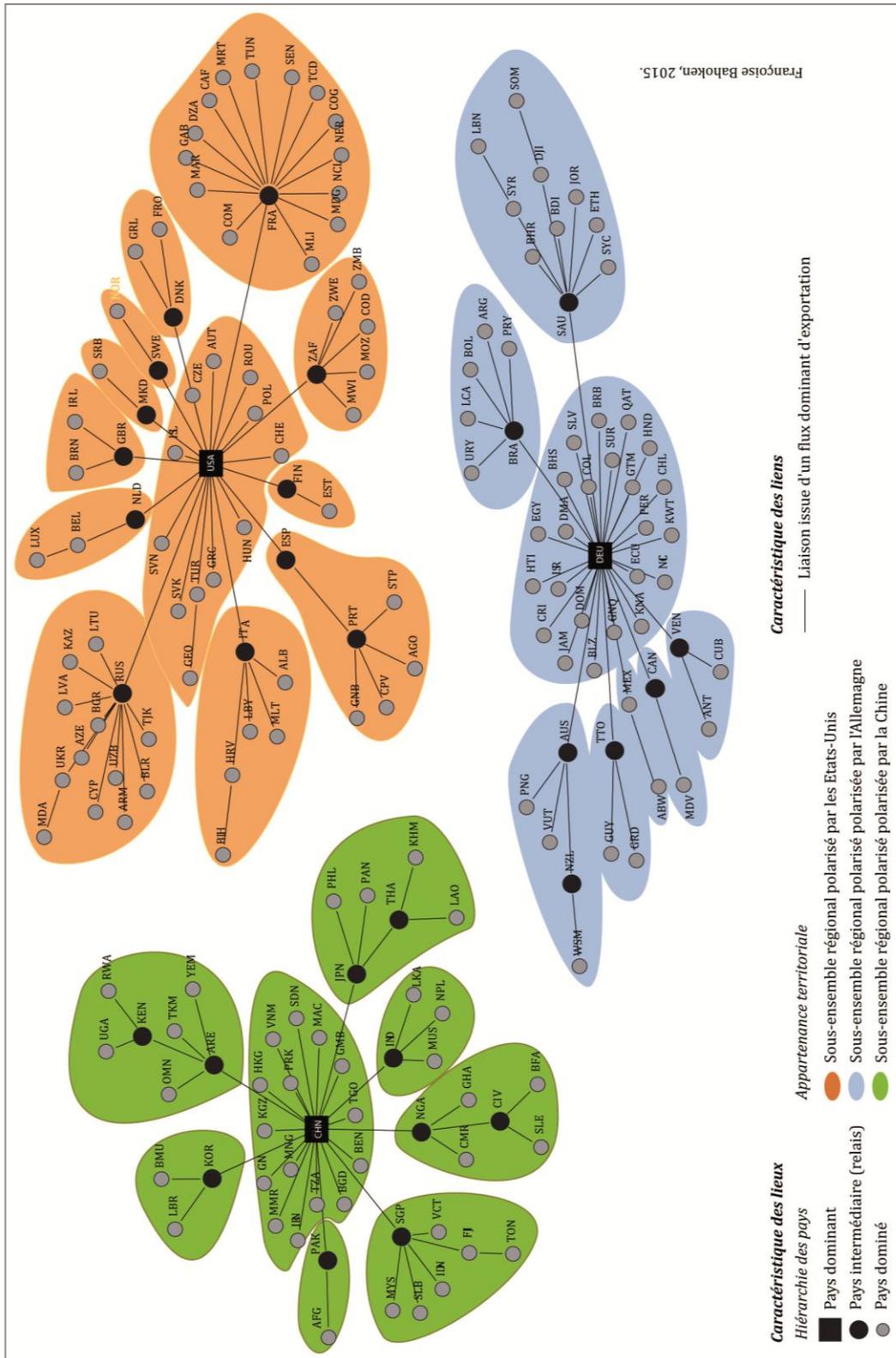
Figure 9-27. Illustration de la variation de la cartographie des *flux dominants* (Fij) versus (Fji)



Deux éléments sont notables sur la Figure 9-27 : la constance de la position de pôle dominant et la variation de certaines positions de relais. Dans cette matrice fictive, le lieu A est dominant dans les deux cas, tandis que le lieu C a une position de relais dans (Fij) et de dominé dans (Fji). Ce résultat n'est pas surprenant : il suscite l'hypothèse du **rôle d'un effet de taille des pays dans la hiérarchie des lieux qui en découle**, qui mérite d'être analysée en fonction du type de flux dont il est question.

La Figure 9-28 présente le graphe des *flux dominants* de (Fji), mettant en évidence la hiérarchie des pays exportateurs.

Figure 9-28. Graphe des régions et sous-régions nodales issues de l'analyse des flux dominants du commerce mondial d'exportation (Fij) de l'OMC (2007)



On distingue à l'échelle mondiale trois grandes régions exportatrices qui sont polarisées par les Etats-Unis (USA), l'Allemagne (DEU) et la Chine (CHN). Ce résultat confirme notre hypothèse quant à la position de la Chine dans ce système mondial de flux commerciaux dominants : elle serait plutôt

importatrice de biens peu onéreux (de ferraille) qu'elle transformerait ensuite en biens de forte valeur qui contribuent à la hisser au sommet de l'arborescence.

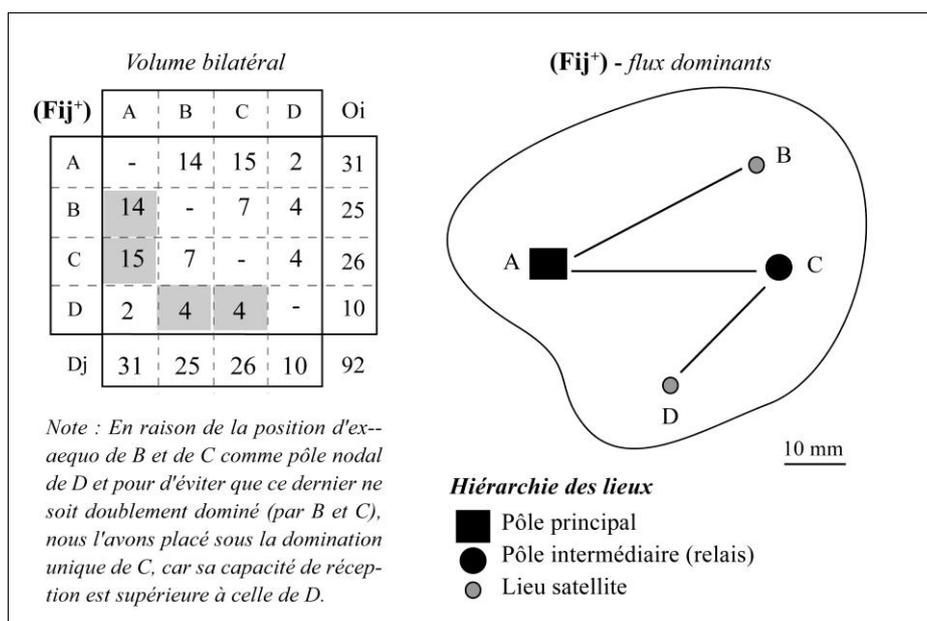
Comme dans le cas de (Fij), il convient toutefois d'analyser avec attention la régionalisation qui découle de cette analyse, en raison des choix méthodologiques sous-jacents concernant les critères de sélection des flux et la méthode de placement des lieux. Sachant que la Figure 9-28 illustre les pays qui émettent beaucoup en valeur, on comprend bien la position de la Chine comme polarisant une région d'exportation, l'« usine du monde » dont on peut considérer qu'elle contribue à générer la richesse mondiale.

Cette difficulté d'interprétation thématique peut être en partie résolue, en mobilisant une matrice symétrique plus évidente à analyser, car elle fait fi de l'orientation des flux.

9.3.3 Le cas évident : la cartographie des flux dominants (Fij⁺)

La cartographie des flux *dominants* d'une matrice symétrique de type (Fij⁺) apparaît évidente pour deux raisons principales : parce que la représentation intègre d'emblée des figurés linéaires non valués et que la symétrisation résout la question de l'orientation des flux. Cette matrice (Fij⁺) présente donc l'avantage de réduire les risques d'interprétation liés à l'utilisation de (Fij) ou de (Fji), mais elle est très artificielle. La Figure 9-29 illustre avec la matrice d'exemple local la cartographie usuelle de ces *flux dominants* (Fij⁺).

Figure 9-29. Illustration de la cartographie de flux dominants (Fij⁺)



On observe sur la Figure 9-29 que la cartographie des *flux dominants* (Fij⁺) correspond en tous points à celle de (Fij), présentée sur la Figure 9-12, page 365. Cependant, elle aurait tout aussi pu apparaître comme une combinaison des deux précédentes (Fij et Fji) présentées sur la Figure 9-27, en présentant plus de relations, si nous n'avions pas appliqué cette règle arbitraire permettant de différencier les lieux *ex-aequo*. Si cette hypothèse apparaît la plus probable, elle mériterait d'être davantage investiguée sur des données empiriques.

Mobiliser (Fij⁺) entraîne un changement sémantique qui n'est pas anodin aux niveaux méthodologiques et thématiques. En effet, la symétrie est une exception en mathématique : dans la réalité, les relations étant généralement asymétriques. Le monde des flux commerciaux n'est pas connu, dans l'ensemble, pour être symétrique même si certaines relations peuvent l'être,

ponctuellement et à titre exceptionnel. Réduire ainsi ce type de matrice engendre donc un risque d'interprétation erronée lié à la création d'exceptions, ou des difficultés imputables à la généralisation du propos. Dans le cas de flux internationaux de marchandises (OMC, 2007), symétriser la matrice conduit à la représentation du volume d'échanges bilatéraux entre des lieux qui deviennent identiques. Il s'agit désormais de pays dont on observe davantage la relation de domination rapportée au volume bilatéral total – le critère [9-2] de la méthode (Nyusten-Dacey) étant par construction égal à la somme des marges ($O_i + D_j$) – et non plus de pays dont on analyse la capacité exportatrice (D_j) ou importatrice (O_i) et inversement. Elle conduit *in fine* à appauvrir l'interprétation.

Si le cas officiel de cartographie des *flux dominants* correspond au résultat obtenu avec la matrice (F_{ij}), il existait donc bien, théoriquement, d'autres cas qui correspondent aux différentes familles de matrices pouvant être déduites de (F_{ij}). Le résultat est par conséquent susceptible de varier en fonction du type de matrice, mais aussi de la manière dont les deux conditions [9.1] et [9.2] sont appliquées, éventuellement d'une troisième permettant de gérer les *ex-aequo*. Cependant, ces critères peuvent être insuffisants pour saisir l'essence des flux concernés, aussi être inadaptés pour certains types de flux. A l'échelle mondiale, par exemple, les flux de marchandises internationaux ne sont pas *symétriques* à l'inverse de certains flux de passagers qui peuvent l'être. Lorsque l'information de flux porte sur la capacité des avions empruntant des lignes aériennes régulières entre deux états touristiques (et non le taux de remplissage), la matrice qui en résulte peut être *symétrique*, ce qui ne sera jamais le cas pour des matrices de flux de marchandises).

C'est parce que ces critères peuvent être inadaptés à la thématique du flux et au niveau d'analyse souhaité que plusieurs auteurs en ont proposé des variantes ou des extensions qui permettent de résoudre des difficultés particulières ou d'affiner l'analyse. La méthode Nyusten-Dacey connaît un vrai succès sur le plan analytique et cartographique (M. Baron et al., 2008 ; N. Cattan, 2004 et G.-P. Torricelli, 1997, l'ont d'ailleurs mis en oeuvre). « *Les nombreuses applications de cet algorithme montrent l'influence de la distance géographique et les effets de barrière (ex : administrative, douanière) sur la formation des régions ainsi obtenues, notamment dans le cas des flux ferroviaires et aériens (Cattan, 1995) et migratoires dans l'étude des pôles d'emploi par l'INSEE* » (Ducruet, 2011 : 5). Cependant, cette méthode est critiquée, car elle conduit à une réduction trop importante de l'information, elle limite aussi l'analyse au *flux majeur*, donc à un seul et unique flux. Les auteurs en sont d'ailleurs conscients. Aussi suggèrent-ils que « [...] *la structure nodale [la hiérarchie des lieux qui en résulte] peut être utilisée pour distinguer des groupes de villes qui ont un lien direct maximum et le rang des autres villes peut être calculé. L'arrière-pays des villes centrales peut également être déterminé, en cartographiant les villes de la structure nodale et en traçant une ligne juste après celles qui sont les plus éloignées de la ville centrale.* » (Nyusten, Dacey, 1961 : 36), définissant ainsi une aire d'association fonctionnelle spécifique à chacun des pôles dominants, fondée sur l'examen des seules relations fortes et directes. « *Elle n'incorpore par les associations indirectes qui pourraient avoir, en théorie, une grande influence dans la détermination des associations fonctionnelles* » (*ibid.*).

Pour pallier à la restriction inhérente à la condition du flux de rang 1, les auteurs suggèrent de l'étendre aux liens indirects, en ordonnant en amont les valeurs de flux. L'association indirecte liée à l'éloignement des lieux par rapport au pôle dominant intègre alors l'hypothèse gravitaire, sous la forme d'une distance géodésique : « [...] *Il est postulé que l'accroissement de l'association indirecte ou de l'influence indirecte [du pôle] diminue avec l'augmentation de la longueur du lien* » (*ibid.*) Autrement dit, plus les villes dominées sont éloignées du pôle dominant et moins elles sont placées sous son influence. Les auteurs en proposent alors plusieurs variantes, parmi lesquelles celles plus récentes de (Rabino et Ocelli, 1997 ; Cattan, 2004 ; Berroir et al., 2007). Si ces propositions sont intéressantes sur le plan théorique, elles n'entraînent pas de modifications majeures sur le plan

méthodologique ou cartographique au sens où elles mobilisent toujours un (voire plusieurs) critères de sélection des valeurs de flux ainsi qu'un paramètre permettant de hiérarchiser les lieux. La construction cartographique sera globalement similaire, fondée sur le principe *liens-lieux*. Ce qui changera donc entre ces variantes, c'est davantage la thématique du flux et l'interprétation qui en sera faite plutôt que la méthode cartographique.

Comme annoncé dans le chapitre 8 (section 8.4), nous proposons dans une dernière section, un ensemble de programmes développés sous RStudio©, permettant d'appliquer les deux conditions de sélection des *flux dominants* de la méthode (Nyusten-Dacey), ainsi que quelques principales variantes proposées par ces auteurs. Elles sont proposées dans *l'Annexe – R. 5. Suggestion de programmes de sélection de flux à cartographier selon un critère local*

Conclusion du chapitre

Les solutions liées à l'application d'un *critère local* fondé sur l'analyse des *flux dominants* de J. D. Nyusten et M. F. Dacey (1961) mobilisent plusieurs familles de méthodes de traitement et de cartographie de flux. C'est pourquoi les résultats qui en découlent sont nombreux. La reconstruction cartographique que nous en avons proposé nous a conduits à identifier les cinq familles de résultats suivants.

La première famille est classique, elle correspond à la cartographie de *flux majeurs* $\text{Max}_i(\text{Fij})$, selon une *logique de flux* classique, entraînant la cartographie d'une matrice (Fij) réduite aux seuls flux de rangs 1. De ce fait, cette solution correspond à une variante des procédures de sélection *globale* associée à une sélection selon la focale de la matrice (les (i) lieux d'origine). La seconde famille de résultats est plus innovante sur le plan conceptuel, car elle introduit un raisonnement sur les lieux, une *logique de lieux*, sur l'analyse de leur attractivité (Dj) en complément de l'analyse préalablement réalisée. Cette logique conduira aussi à celle de leur hiérarchisation. Cette seconde famille conduit à une double restriction de la matrice de flux, respectivement analytique et graphique : la première transforme les *flux majeurs* en *flux dominants* tandis que la seconde substitue des liens aux flux. On montre ainsi que les auteurs adoptent en fait une logique de représentation *liens-lieux* plutôt que de flux. La troisième famille de solution est d'ordre sémiologique et conceptuel, elle porte sur la symbolisation de la hiérarchisation des lieux de ce système de *flux dominants* par le truchement de la *Forme* à la place de la *Taille*, plus logique. Cette posture introduit alors un raisonnement d'ordre *qualitatif* sur l'attractivité des lieux (en éludant la *Taille*) et non plus uniquement *quantitatif*, comme dans la *logique de flux* classique. La quatrième série de résultats est liée à la formation de régions nodales, polarisées par des lieux qui en dominent d'autres. Le processus de régionalisation étant hiérarchique, nous avons montré qu'il conduisait à une décomposition en plusieurs niveaux caractérisés a *minima* par des lieux *dominants*, *dominés* ou intermédiaires. Cette hiérarchisation correspond à une cinquième famille de résultats qui identifie des sous-bassins au sein de la partition principale. Ces deux dernières familles correspondent, en fait, à une variante du cas des partitions de l'espace des flux en (Aij) présenté dans le chapitre 7 (voir section 7.3) à la différence près que la présente approche est fonctionnelle et non structurelle : elle est fondée sur l'exploitation des valeurs des liens et non des lieux. De ce fait, les régions et sous-régions ainsi formées peuvent être interprétées comme formant le périmètre de l'aire d'influence des pôles nodaux.

Cette variété de résultats souligne l'importance de la méthode (Nyusten-Dacey), dans un contexte de nécessaire réduction des problèmes de complexité graphique et analytique de la carte de flux. Plusieurs points méritent d'être soulignés.

La méthode met en évidence la complémentarité analytique du graphe et de la carte de flux. Elle apporte une bonne réponse à trois principaux problèmes de l'analyse cartographique de flux. Elle résout les problèmes de :

- *surcharge graphique*, par la suppression de la *largeur* des liens ;
- de *couverture spatiale*, en suggérant une procédure de régionalisation entraînant une représentation en aplats rendant celle des liens facultatifs ;
- de *taille des lieux* en analysant leur attractivité selon un raisonnement *qualitatif* conduisant à mobiliser la *Forme* à la place de la *Taille* ;
- de *convergence des flux*, en supprimant l'orientation et par conséquent le tracé des têtes de flèches.

La méthode s'avère très efficace pour sélectionner les valeurs de flux sur lesquelles une analyse plus poussée pourra être réalisée. En effet, en poursuivant la logique de formation des régions, il est ensuite possible de discuter de la hiérarchisation des lieux envisageable et des partitions obtenues, selon

plusieurs directions : du point de vue des lieux d'origine (F_{ij}) ; des lieux de destination (F_{ji}) ; du volume bilatéral de flux (F_{ij}^+) ou encore du mouvement (M_{ij}^+).

Au niveau cartographique de l'implantation spatiale, la méthode se révèle être la plus riche : elle mobilise le cœur de la matrice (F_{ij}) mais également au moins une marge (D_j) , par conséquent à la fois les liens et les nœuds sous la forme respective de traits, de points et de surfaces pour décrire l'attractivité des lieux et leur aire d'influence.

Enfin, sur le plan méthodologique, la cartographie des *flux dominants* mobilise un raisonnement mixte impliquant, d'une part, la *logique de flux* (ou de *mouvement*) classique, dans le cadre d'un *raisonnement quantitatif* qui mobilise la *Taille* pour représenter le *flux majeur* et, d'autre part, une *logique de lieux discrète*, correspondant à un *raisonnement qualitatif* utilisant la *Forme* et la *Couleur*.

Toute cette richesse sémiologique n'est assurément pas sans effets au niveau analytique, sur la compréhension de l'information de flux représentée. Cette question n'est pas sans importance puisque ici, comme dans d'autres domaines de la visualisation de données, la validité de la traduction cartographique d'un processus réside dans l'interprétation qui en faite.

CONCLUSION DE LA DEUXIEME PARTIE

« *La relation entre villes n'est pas le seul système qui puisse être défini par un réseau de points et de lignes. Les nations ou les États peuvent être considérés comme des points, les migrations ou les flux de marchandises comme des lignes. L'étape importante, lorsque l'on mobilise cette figure abstraite du graphe, réside dans la recherche d'une signification plausible des points et des lignes, de préférence au regard des processus du monde réel.* »

John D. Nyusten et Michael F. Dacey (1961).

L'objectif de cette deuxième partie était de proposer des solutions visant à réduire la complexité graphique de la cartographie d'une matrice de flux. Pour cela, nous avons fait le choix d'agir, en priorité, sur la réduction voire la suppression de certaines des contraintes graphiques identifiées dans la première partie. Notre attention a en effet été mobilisée sur les effets de *surcharge graphique*, de *couverture spatiale*, concourant à l'*effet spaghetti*, ainsi que de *disposition* des figurés ; toutes impactant directement la perception des motifs de flux représentés sur la carte. Ces contraintes graphiques étant le résultat de méthodes de traitement des données numériques et cartographiques, mises en œuvre en amont du processus de représentation, l'investigation sur les solutions envisageables intéresse à la fois l'apparence de l'objet carte de flux (les aspects visuels et esthétiques), que le processus de cartographie de la matrice de flux proprement dit (le traitement des données et leur représentation).

Autrement dit, la réponse apportée à l'objectif de cette seconde partie concerne d'une part, l'objet carte et d'autre part, son processus d'élaboration. Dans le premier cas, il s'agit de réduire la complexité visuelle, ou perçue par l'œil humain, qui est liée à l'observation de l'information représentée, au ressenti sensible produit face à la figure, selon qu'elle est claire ou non, que le message est lisible ou brouillé... Dans le second cas, en lien avec le premier, il convient de réduire la complexité analytique en s'assurant que la figure est lisible non seulement sur le plan graphique mais aussi au niveau cognitif, en décrivant des motifs aisément compréhensibles par un observateur statique. La réduction de cette complexité ayant donc pour objectif général de rendre intelligible l'information de flux cartographiée. Enfin, l'approche mise en œuvre de simplification du processus de cartographie de flux ayant une vocation générale, généralisable et reproductible, elle a également concerné les aspects liés à la mise en œuvre pratique des éventuelles solutions proposées

Alors que les solutions apportées jusque là était essentiellement liées au traitement des données numériques ou cartographiques, nous avons montré (chapitre 5) qu'il était possible d'agir au seul niveau graphique, sur la sémiologie des flux pour réduire le problème de la complexité graphique de la carte de flux. La mise en œuvre de la *Transparence* sur des versions décomposées (Chapitre 3) de la matrice de flux, en l'occurrence sur (Lij) , (Fij^+) et (Fij^-) , associée à une utilisation inédite de la couleur, contribue indéniablement à renouveler l'image des flux. Si l'apparence des cartes qui en résulte est modifiée, il est important de noter que le processus de construction cartographique ne l'est pas ; sa mise en œuvre pratique est d'ailleurs simplifiée. En effet, pour (Fij^-) par exemple, deux innovations sont notables : d'une part, l'utilisation de bandes non orientées pour représenter des

données orientées, grâce au transfert de l'opposition colorée au niveau des bandes ; d'autre part, dans le dédoublement du nombre de figurés à tracer qui résout le problème du choix de la partie de la matrice à cartographier (les valeurs positives ou les valeurs négatives) et cela, sans opérer de sélection des valeurs.

Les solutions proposées ont également un impact fort sur la représentation de la dialectique *liens-lieux*, résolvant ainsi le problème de la représentation de *l'effet de convergence* des flux. En appliquant les solutions sémiologiques sur des données réelles, nous avons observé qu'il n'était plus nécessaire de représenter conjointement les flux et leur effet au niveau des lieux, sans nuire à la perception de l'information. Toutefois, ce gain ne signifie pas que *l'effet de taille des lieux* soit gommé ou moins perceptible. Il l'est tout autant, sous une autre apparence graphique (voir Figures 5-11 et suivantes, à partir de la page 233). L'intérêt thématique de cette méthode, qui tient dans la hiérarchisation apparente des flux (F_{ij}) apportée par le recours à la *Transparence*, est de conduire à deux hypothèses fortes : d'une part, **l'existence d'une forme d'attractivité des lieux représentée par l'effet de polarité des bandes orientées, en opposition colorée** pour traduire leur orientation et cela, sans recourir à la flèche. D'autre part, **l'existence de différentes classes de longueurs de bandes dirigées vers ces lieux attractifs**, ce qui correspond, dans la réalité, à différentes classes de distances perçues des flux ou bien à différents types de mobilité spatiale.

Ces propositions sémiologiques permettent de résoudre nombre de problèmes graphiques, aussi d'en révéler d'autres, tels qu'un *effet de profondeur*. Elles ont aussi ouvert la voie vers la prise en compte de l'espace dans le processus de construction cartographique. Tenir compte de la géographie des positions (x, y) qui caractérise l'espacement des lieux d'origine et de destination, dans un système de projection donné, nous a conduits à mobiliser la notion *distance parcourue par le flux* sous différentes formes. Si son acception est courante en analyse spatiale, comme variable spatiale d'éloignement, nous avons montré (chapitre 7) l'intérêt d'une approche discrète fondée sur les notions de *voisinage* spatial ainsi que d'appartenance territoriale commune. L'intérêt théorique et analytique d'un tel raisonnement sur l'espace des flux par le truchement de différentes formes d'espacement des lieux dans le processus de cartographie est indéniable et non seulement lors du traitement des données, par exemple, pour la mise en évidence d'effets de barrières. Cet intérêt est également pratique, au niveau graphique. La réduction de la matrice de flux apportée par l'application d'un critère de sélection fondée sur la contiguïté de voisinage ($C_{ij}=1$) ou (CK_{ij}) entraîne automatiquement l'économie de procédures additionnelles de sélection de l'information à représenter ou de réduction de la matrice (chapitres 8 et 9).

Cette seconde partie concerne également les méthodes usuelles de cartographie de flux fondées sur le principe de sélection de l'information à représenter. Certaines d'entre elles, pourtant très répandues (comme celle de la sélection des 50 plus gros flux), probablement en raison de la facilité apparente de leur mise en œuvre, sont étendues dans le cadre d'un raisonnement matriciel, et plus robuste sur le plan statistique. On montre ainsi que la carte portant sur une sélection de valeurs de flux peut être réalisée de différentes manières tout en renouvelant l'image de ces flux. Deux niveaux de raisonnement ont été proposés dans l'optique de la cartographie d'une partie réduite de la matrice. Le premier niveau agit au niveau global de la matrice (chapitre 8), la mobilisation d'un critère graphique nous conduit à proposer le recours à un (double) critère de sélection qui agit sur la part d'information de flux (F_{ij}) à représenter en fonction de la part de liens (L_{ij}), soit sur une courbe de Lorenz dont le degré de concentration des valeurs correspond à l'indice de Gini. Le second niveau agit localement, c'est-à-dire sur les localités (ou lignes et colonnes de la matrice) en focalisant d'avantale l'analyse sur les lieux de destination (ou d'origine) que sur les liens. La sélection de l'information de flux locale est proposée dans le cadre d'analyse de *flux dominants* développé par J-D. Nyusten et M-F. Dacey (1961). La reconstruction proposée de la méthode originale met en évidence un raisonnement qualitatif qui,

parce qu'il privilégie la qualité des lieux au regard de la quantité des liens (seules les liaisons issues de *flux dominants* sont représentés sur la carte) introduit une *logique de lieux* discrète (pour représenter la hiérarchie des lieux) aux cotés de la traditionnelle *logique de flux* (ou plutôt *logique de liens*, les auteurs ne représentent à aucun moment la valeur des flux). La richesse d'un raisonnement focalisé sur l'ensemble des parties de la matrice (les liens et les lieux) trouve sans doute ici sa plus belle expression.

L'intérêt d'un raisonnement mené en fonction du type de matrice et du type de flux (Fij) est avéré, et validé par des procédures variées de sélection des flux. On valide ainsi l'hypothèse, présentée dans l'introduction générale, selon laquelle **la réduction des difficultés d'ordre méthodologique de la cartographie de flux est inhérente à une bonne connaissance du maniement des matrices de flux, aussi de la thématique des flux dont il est question** sauf à s'associer à un expert de ces derniers pour l'interprétation des résultats.

Nous avons en effet montré que ces questions méthodologiques interviennent en réalité dans un second temps de l'analyse, sans pour autant être secondaires. La raison tient à l'invariance de la construction (carto)graphique du flux, que nous avons qualifiée de *logique de flux*, quel quel soit le type de matrice (la *logique de flux* consiste à agir sur la seule dimension de *largeur* du figuré). Cette dernière ne faisant pas état de la composante spatiale, nous avons introduit une *logique de mouvements* permettant de mobiliser l'espace des positions dans la définition de l'information, par le recours à une acception de la distance parcourue. On démontre ainsi en utilisant un cadre d'analyse sémiotique de la construction du figuré qui illustre le flux que la carte de flux (Fij), tout comme le graphe spatialisé, sont similaires sur le plan cartographique et sur le plan sémantique en ce sens qu'ils font fi de l'espace. Ils le sont également avec le mouvement (Mij), mais du seul point de vue cartographique, ce qui prouve l'invariance de la solution classique. Par conséquent, ce qui va distinguer une cartographie de flux et une cartographie de mouvement ce n'est pas tant la manière dont le flux est représenté – quelle que soit l'échelle, dans un système de projection cartographique donné – que l'interprétation qui est faite, autrement dit la sémantique.

Nous avons ainsi pu apporter un éclaircissement théorique quant à la signification de la carte de flux, grâce à son repositionnement dans le contexte de l'analyse géographique, au sein d'un territoire façonné par des sociétés où les positions relatives des lieux les uns par rapport aux autres n'est pas aléatoire (à l'inverse du graphe). L'investigation menée au niveau cognitif de la perception de l'information représentée selon une logique de flux par rapport à une logique de mouvement nous a permis d'aboutir à la conclusion suivante. Dans un contexte où la projection fait sens, en raison du formalisme du trait droit (nécessaire pour ne pas altérer l'information), ajouté à l'impossibilité de traduire avec exactitude la distance-temps parcourue par le flux depuis la sphère vers la carte, la cartographie d'une matrice de flux spatialisés ne saurait être qu'une approximation de celle d'un mouvement.

CONCLUSION GENERALE

« Mouvement, mouvement, mouvement... cela parce que la plupart des changements en Géographie sont dus au mouvement ».

Waldo Rudolf Tobler.

L'objectif général de cette thèse était de proposer des solutions générales et opérationnelles permettant de réduire la complexité visuelle, *a fortiori* analytique, qui résulte de la cartographie d'une matrice de flux dense. Pour ce faire, nous avons mené un raisonnement en deux temps, correspondant aux deux grandes parties.

Dans la première partie, nous avons proposé un état de l'art des différentes méthodes de (carto)graphie des matrices de flux. Celui là nous permettant, d'une part, de formaliser la pratique usuelle et les différents problèmes et contraintes qui concourent à la complexité de la cartographie de flux : de la mise en œuvre du processus à l'image qui en résulte. D'autre part, d'identifier les contraintes graphiques et les difficultés d'ordre méthodologique inhérentes à la construction des différentes possibilités de cartographie de flux.

Dans la seconde partie, nous avons proposé un ensemble de solutions à une série de problèmes identifiés dans la première partie. Leur choix s'inscrit dans une volonté de cohérence des propositions quant à l'objectif à atteindre : la réduction de la complexité de la figure. De même, les solutions proposées sont cohérentes sur le plan des méthodes à mobiliser. Elles revêtent, en effet, un caractère général et sont généralisables ; elles ont été présentées sur une matrice fictive récurrente qui sert d'exemple de démonstration, puis sont validées sur des matrices variées portant sur différents types de flux migratoires ou commerciaux qui s'expriment à différentes échelles géographiques.

Notre recherche de solutions liées à la difficulté de la cartographie des matrices de flux s'est focalisée sur le sens de l'information représentée, sa construction, sa perception et la compréhension que sa cartographie apporte. La résolution des contraintes graphiques et des difficultés d'ordre méthodologique identifiées a consisté en des actions variées que nous avons menées à différents moments du processus cartographique. Elle a donné lieu à plusieurs possibilités de résolution, à des solutions que nous avons examinées pour des raisons de complétude théorique, de démonstration et de validation empirique à différentes échelles géographiques. Si nous avons également proposé une formalisation générale de l'approche usuelle de cartographie de flux, il est également possible de mentionner nos apports.

Une série de cartes valant mieux qu'un long discours, nous proposons de synthétiser visuellement ces apports à travers trois planches que nous commenterons brièvement.

Principaux apports de la thèse

Une conclusion étant, en fait, toujours l'occasion d'expliquer et de mettre en avant les principaux apports du travail réalisé, elle donne généralement lieu à une liste. Pour présenter les nôtres, nous n'allons pas les lister (cela est déjà fait dans les conclusions de chapitre). Nous allons plutôt avoir une approche, certes peu conventionnelle, mais qui correspond parfaitement à une mise en scène cartographique. Après tout, ne s'agit-il pas d'une thèse en cartographie ?!

La mise en scène cartographique que nous proposons consiste donc à sélectionner un type de figure que nous considérons le plus représentatif de nos apports. Les exemples d'application récurrents dans la thèse étant dédiés à des flux migratoires – des migrations internes à l'ex-Tchécoslovaquie (1990) et des navettes locales en Isère, France (2010) –, nous choisissons pour conclure de revenir sur des figures de validation empirique portant toutes sur des flux commerciaux internationaux et mondiaux. Les trois planches suivantes permettent ainsi de se focaliser, directement et graphiquement, sur des exemples de trois familles d'apports.

Les apports théoriques et conceptuels, présentés sur la **Planche 1**, sont issus des chapitres 3 et 6. Ils s'inscrivent dans le cadre d'une formalisation de la pratique usuelle de cartographie de flux que nous avons qualifiée de *logique de flux*. Elle mobilise la *Taille* et le formalisme du trait droit pour symboliser la valeur du *flux* dans un plan quelconque, justifiant ainsi le recours indépendant au *graphe* ou à la *carte de flux*. Cette *logique de flux* s'oppose à une *logique de mouvement* qui, si elle mobilise également la *Taille*, place à l'inverse l'analyse dans un territoire, justifiant en ce sens le tracé sinueux du flux inscrit dans un espace : sa perception correspondant à peu près au cheminement spatio-temporel effectivement réalisé. Ces deux *logiques de flux* ou *de mouvement* sont inscrites dans un raisonnement exclusivement *quantitatif* : les mesures correspondantes portent sur des effectifs. Elles s'opposent de fait au *raisonnement qualitatif* qui mobilise plutôt les modalités d'un caractère, des caractéristiques qui conduisent à mobiliser sur le plan sémiologique des variations de *Forme*.

Le *raisonnement qualitatif* correspond à la *logique liens-lieux* présentée sur la **Planche 2**, issue des apports méthodologiques du chapitre 9. La « reconstruction » cartographique de la méthode (Nyusten, Dacey, 1961) a montré l'importance accordée par les auteurs aux caractéristiques qualitative des lieux (*Forme*) – selon qu'ils sont dominants ou non – plutôt qu'à celles des flux (*Taille*), représentés respectivement par des *points* et des *liens* non valués. C'est aussi pourquoi nous considérons que cette méthode de sélection des flux est locale. Elle est, en effet, focalisée *i*) sur les localités (les marges de la matrice et non sur leur cœur) dans l'objectif de représenter la mise en relation hiérarchique des lieux issue d'une analyse des *flux nodaux* (dits *dominants*) et *ii*) sur l'identification des *régions nodales* (symbolisées par des polygones) qui en découlent plutôt que les interactions territoriales en elles-mêmes (qui l'auraient été par des bandes dans la *logique de flux* classique). Nous avons volontairement opté ici pour une représentation graphique plutôt que cartographique, afin de souligner l'efficacité des graphes non spatiaux pour certains sujets tels que la mise en évidence d'un arbre hiérarchique.

Les apports illustrés par la **Planche 3**, répondent à un tout autre raisonnement puisque ce dernier est visuel. En ne portant que sur des critères d'ordre sémiologique, liés à l'apparence des images, ils s'inscrivent dans un *raisonnement graphique* qui agit sur l'esthétique de l'image, sur le tracé des flux et sur leur sémiologie. L'action est menée principalement sur la couleur mobilisée en tant que teinte (noir/blanc) et variable visuelle (*Couleur*). Le recours au critère de *Transparence* suggéré par (Mac Eachren, 1995) conduit à la possibilité inédite de représenter toute *l'information disponible* dans une matrice dense. De même que le transfert du principe de *l'opposition colorée* au niveau des liens permet de renouveler la cartographie des matrices de flux *antisymétriques*. Ce dernier apport est sans doute l'un des plus originaux de notre travail : il permet en effet de faire surgir une vision originale des pôles d'émission et d'attraction qui reprend la logique des champs vectoriels chers à W. Tobler.

Planche 1. Apports théoriques et conceptuels : l'exemple de la construction de la carte de flux (ou de mouvement)

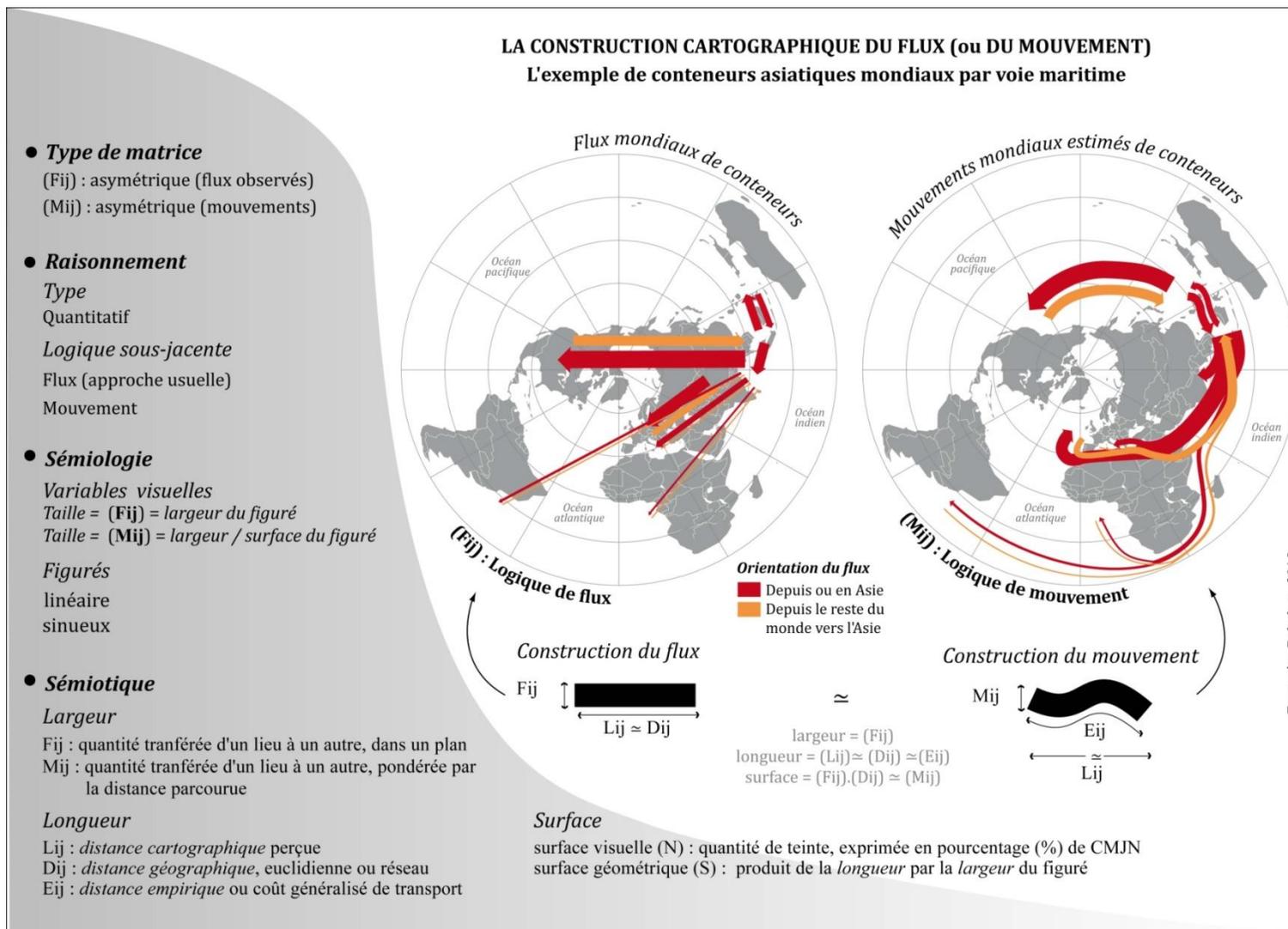


Planche 2. Apports méthodologiques : l'exemple du critère de sélection local de l'information de flux à cartographier

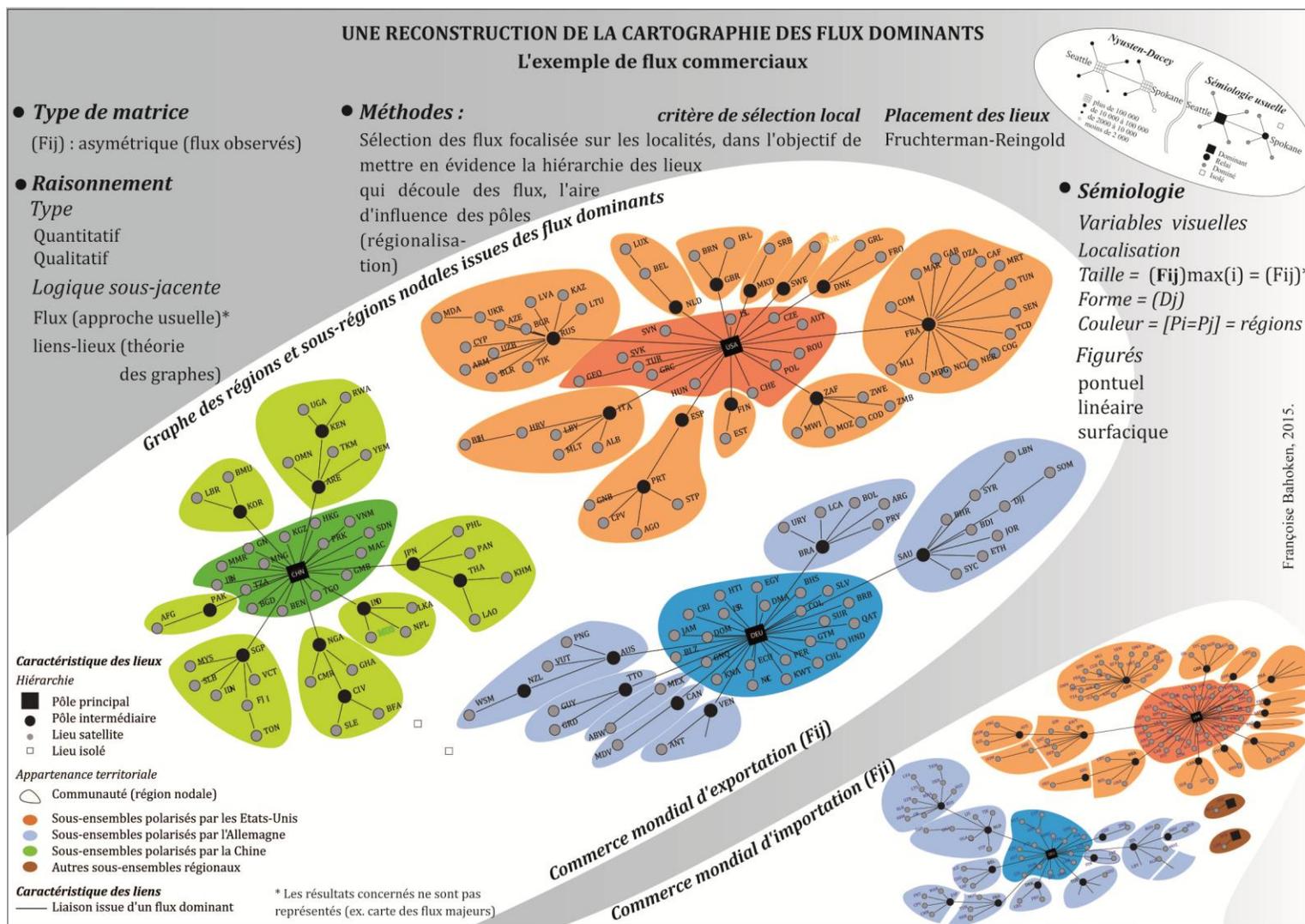
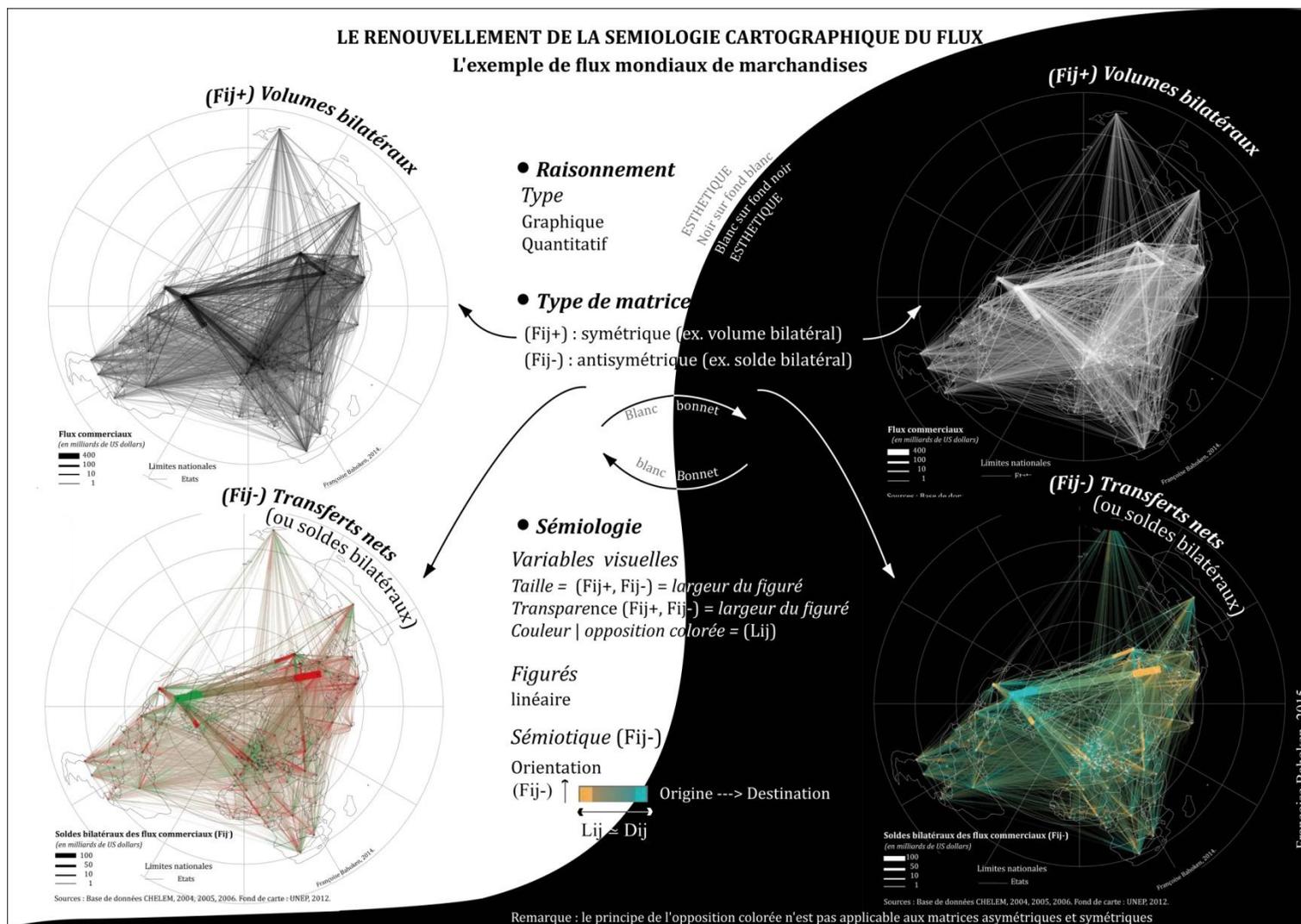


Planche 3. Apports sémiologiques : le renouvellement de la variable visuelle *Couleur*



Perspectives

Les solutions proposées dans cette thèse sont générales, elles ont été mises en œuvre sur une matrice unique (Fij), à l'aide de méthodes d'analyses descriptives s'intéressant au fond de la carte et à sa représentation (carto)graphique ; elles n'ont pas mobilisé de méthodes conduisant à modifier le fond de carte, ni des matrices multiples ou complexes. Dès lors, plusieurs pistes sérieuses d'approfondissement ou de poursuite des travaux initiés peuvent être engagées.

Une première piste consisterait à **valider empiriquement la conclusion selon laquelle la cartographie de flux serait en réalité une représentation approximative du mouvement**. Pour cela, il faudrait introduire le rôle de l'utilisateur (l'observateur), en réalisant une enquête par questionnaire permettant la collecte d'informations sur les mouvements oculaires et les motifs perçus par l'œil lors de la visualisation de cartes de flux et de mouvement. Cette enquête pourrait être menée dans le cadre d'une collaboration avec des chercheurs des sciences cognitives, de l'information et de la communication, soit plus généralement en sémiotique.

Une seconde piste concernerait **la résolution d'effets non pris en compte**, tels que les effets généraux liés à la *position des lieux* qui pourraient être examinés dans le cadre de solutions liées à des procédures de transformations cartographiques. Trois familles de solutions, intéressant les aspects graphiques du tracé et de perception de l'information, pourraient être envisagées : la première examinerait la position des lieux d'origine et de destination (dans le formalisme du trait droit) ; la seconde concernerait le tracé des liens (dans le cas d'une projection cartographique fixe) et la troisième mobiliserait des méthodes d'agrégation (de *clustering*) conduisant à un changement de résolution de l'information.

Une troisième direction porterait sur **le développement d'outils susceptibles de faciliter le processus de cartographie de données provenant de différents types de matrices (de flux (Fij) ou de mouvements (Mij)) et leur appropriation par le plus grand nombre**. Pour cela, il serait possible de poursuivre et finaliser les développements engagés dans cette thèse (disponibles en annexe) puis de les implémenter dans une application libre qui pourrait être disponible en ligne, sous une forme interactive. En effet, l'urgence en la matière est de proposer des outils innovants sur le plan méthodologique, efficaces sur le plan de la construction cartographique mais également sur les plans techniques et technologiques dans une configuration facilitant leur appropriation. Quant au choix de la technologie, il est nécessaire de recourir à des solutions simples, générales, reproductibles et surtout ouvertes. C'est pourquoi ces développements pourraient être réalisés sur la plateforme Rstudio et prendraient, dans un premier temps, la forme de *package* avant celle d'application (RShiny, par exemple). Nous pouvons mentionner, à titre d'exemples, des travaux en cours de développement : le package *Cartography*²¹⁶ à peine déposé par T. Giraud (2015) ; les applications : *mapshaper.org*²¹⁷ (2013) développée et maintenue par M. Bloch ou encore *Trajpop*²¹⁸ (2013), développée et maintenue par R. Cura, dans le cadre du projet européen GeoDiverCity²¹⁹.

Une quatrième direction porterait **sur le développement de solutions spécifiques au cas des matrices complexes : catégorielles et/ou temporelles et/ou multi niveaux géographiques**. En effet,

²¹⁶ La vignette (ou document descriptif) de ce package est disponible sur le site du CRAN, à l'adresse suivante : <https://cran.r-project.org/web/packages/cartography/vignettes/cartography.html> [lien vérifié le 13/10/2015].

²¹⁷ Cette application correspond à une actualisation pour internet du programme éponyme que l'auteur avait développé en 2006, à l'Université du Wisconsin ; elle a été mise hors ligne le 27 janvier dernier, ce qui témoigne de l'intérêt ces nouvelles technologies liées à l'internet. Plus d'informations est disponible sur le compte *GitHub* de l'auteur : <https://github.com/mbloch/mapshaper/commit/dca299bfc9da39e20d70f85ec50799df59c1765b>

²¹⁸ L'application *Trajpop* est accessible à l'adresse suivante <http://trajpop.parisgeo.cnrs.fr/> [lien vérifié le 05/05/2015].

²¹⁹ Le site internet de ce projet est accessible à l'adresse suivante : <http://geodivercity.parisgeo.cnrs.fr/blog/> [lien vérifié le 02/08/2015].

il nous apparaît intéressant de réaliser des analyses des changements inter matriciels, à partir de l'exploitation des différentes parties complémentaires d'une matrice unique portant sur le total de flux (Fij). Cette analyse pourrait passer par la comparaison d'un jeu de plusieurs matrices qui seraient indépendantes et complémentaires, car disponibles sur la même zone d'étude. Ces hypermatrices rajouteraient une troisième dimension (temporelle ou catégorielle) aux deux dimensions spatiales déjà disponibles que sont l'origine et la destination.

Trois cas pourraient être envisagés : celui où (Fij) serait décomposable en plusieurs catégories d'individus (Fij) \rightarrow (Fijk) avec (k : k_a, k_b, ..., k_n) et/ou diachronique (Fij) \rightarrow (Fijt) avec (t : t₁, t₂, ..., t_n). Une hypothèse provisoire serait que la mobilisation de ces matrices complexes – dans un objectif de cartographie de flux – conduirait à la réalisation de *cartes de synthèse de flux (ou de mouvements)* qui superposeraient éventuellement différents thèmes, à la manière du croquis de synthèse. L'objectif ne serait pas tant de réaliser des cartes différentes dans un but comparatif mais plutôt d'examiner le résultat de la combinaison de plusieurs variables (de plusieurs matrices) sur la perception des motifs représentés, en limitant la perte d'information.

Si plusieurs approches de l'analyse de données coexistent, elles n'ont pas toujours été examinées d'un point de vue cartographique... et c'est là qu'on s'aperçoit que de nombreuses hypothèses demeurent envisageables dans le contexte de l'analyse cartographique des matrices de flux denses. Toutes ces pistes encore inexplorées me conduisent à conclure à la manière de F. Kafka (1984)²²⁰ en disant « [...] *qu'au début de mon procès je voulais le finir et qu'à la fin je ne voulais que le recommencer ?* ».

²²⁰ Cette citation hors champs de la thèse est extraite de : Kafka F., 1984, *Le Procès*, vol. 1, Chapitre 10, p. 464.

BIBLIOGRAPHIE

Cette bibliographie recense les articles et ouvrages cités dans le texte, les documents cartographiques remarquables présentés, les références techniques mentionnées ainsi que mes références personnelles réalisées en lien avec le sujet.

Articles et ouvrages

Cette section liste les références mentionnées dans le texte, par ordre alphabétique du nom du premier auteur puis par ordre chronologique croissant.

Andrienko, N., Andrienko, G. (2004), Interactive visual tools to explore spatio-temporal variation : Coastabile, M.F. (Ed.), *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI'2004)*, Gallipoli, Italy, May 25-28, published in ACM Press, pp.417-420, ISBN 1-58113-867-9, [En ligne] <http://geoanalytics.net/and/papers/avi04.pdf> [lien vérifié le 02/08/2015]

Andrienko, G., Andrienko, N. (2007), Designing Visual Analytics Methods for Massive Collections of Movement Data, *Cartographica*, vol. 42, n°2, pp. 117-138.

Andrienko, G., Andrienko, N. (2008), Spatio-temporal aggregation for visual analysis of movements, *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, pp. 51-58, 2008. [En ligne] <http://geoanalytics.net/and/papers/vast08.pdf> [lien vérifié le 02/08/2015]

Antoni, J.-P., Klein, O., Moisy, S. (2012), Une discrétisation temporelle. Une méthode de structuration des données pour la cartographie dynamique, *Cartes et géomatique*, n°213, pp. 27-31.

Banos, A., (2001) A propos de l'analyse spatiale exploratoire des données, *Cybergeo : European Journal of Geograph*, rubrique *Systèmes, Modélisation, Géostatistiques*, article 197, [En ligne] <http://cybergeo.revues.org/4056> [lien vérifié le 02/08/2015]

Baron, M., Eckert, D., Jegou, J. (2012), Peut-on démêler l'écheveau mondial des collaborations scientifiques ? *M@ppemonde*, Rubrique Internet, n°102, [en ligne] <http://mappemonde.mgm.fr/num30/internet/int11201.html> [lien vérifié le 02/08/2015]

Baron, M., Le Neindre, C., Zanin, C. (2008), Mettre en cartes les universités du Bassin parisien, *M@ppemonde*, n°92, [en ligne] <http://mappemonde.mgm.fr/num20/articles/art08401.html> [lien vérifié le 11/10/2015]

Barthelemy, M. (2010), Spatial networks, *Physics Reports*, vol. 499, Issue 1-3, pp. 1-101.

Bavoux, J.-J., Beaucire, F., Chapelon, L., Zembri, P. (2005), *Géographie des transports*, Armand Colin, Coll. U, 231 p.

Beauguitte, L. (2010), *Graphes, réseaux, réseaux sociaux : vocabulaire et notation*, Méthodes, Synthèses du Groupe f.m.r., 7 p., [en ligne] https://halshs.archives-ouvertes.fr/file/index/docid/541898/filename/fmr1_vocabulaire_notation.pdf [lien vérifié le 02/08/2015]

Beguïn, M., Pumain, D. (1994, 2003), *La représentation des données géographiques. Statistique et cartographie*, Armand Colin, Coll. Cours Géographie, 2^{de} édition, 192 p.

Berry, B. (1964), Approaches to regional analysis: a synthesis, *Annals of the Association of American Geographers*, vol. 54, pp. 2-11.

-
- Bertin, J. (1967), *Sémiologie graphique, Les diagrammes - Les réseaux - Les cartes*, Editions Gauthier-Villiar.
- Bertin, J. (1969), Graphique et mathématique : généralisation du traitement graphique des informations, *Annales, Economie, Sociétés, Civilisation*, vol. 24, n°1, pp. 70-101. [En ligne] http://www.persee.fr/web/revues/home/prescript/article/ahess_0395-2649_1969_num_24_1_422034 [lien vérifié le 02/08/2015]
- Bertin, J. (1973), *Sémiologie graphique*, Paris, Mouton-Gauthier-Villars, 2ème édition, Les réimpressions de l'EHESS (1998), 431 p.
- Blin, E., Bord J.-P. (1993), *Initiation géo-graphique, ou comment visualiser son information*, Sedes, deuxième édition, 284 p.
- Bonin, S. (1975), *Initiation à la Graphique*, EPI éditeurs, 168 p.
- Bopda, A. (1997, 2003), *Yaoundé et le défi camerounais de l'intégration : à quoi sert une capitale d'Afrique tropicale ?* Thèse de doctorat en Géographie de l'Université Paris 1, CNRS éditions, Coll. Espaces et milieux, 422 p.
- Bopda, A., Grasland, C., Poulain, M. (2000), Evaluation comparative de l'influence des limites linguistiques sur les comportements migratoires : applications aux cas de la Belgique, du Cameroun et de la Tchécoslovaquie, in : « Régimes démographiques et territoires : les frontières en question », *Actes du colloque de l'AIDELF Les frontières en question*, 1988, La Rochelle, AIDELF, 20 p.
- Börner, K. (2010), *Atlas of Sciences. Visualizing what we know*. Massachusetts : The MIT Press, 254 p.
- Boyandin, I., Bertini, E., Lalanne, D). (2010), Using Flow Maps to Explore Migrations Over Time, *Proceedings of the International Conference on Geographical systems AGILE' 2010*, Guimarães, Portugal, 4 p.
- Boyandin, I., Bertini, E., Bak, P., Lalanne, D. (2011), Flowstrates : An Approach for Visual Exploration of Temporal Origin-Destination Data, *Proceedings of the Computer Graphics Forum*, vol. 30, n°3, pp. 971–980, 2011.
- Bras (Le) H. (2000), Les lois du Mouvement, Chapitre IV, in : Le Bras, H., *Essai de géométrie sociale*, Odile Jacob, pp. 107-126.
- Breukelman J., Brink G., De Jong T., Floor H. (2009), Selecting Flow Data, in : *Manual. Flowmap software v.3.2*, Faculty of Geographical Sciences, Utrecht University, the Netherlands, pp. 42-44.
- Brewer, C., Harrower, M. (2009), *Colorbrewer: Color advice for maps*. [En ligne] <http://colorbrewer2.org/> [lien vérifié le 02/08/2015].
- Brinton, W.C. & ass. (eds.) (1939), *Graphic Presentations*, Brinton, W. C. & Associates, New York, 505 p. (Cartes de flux : pp. 216-230 ; graphes : pp. 73-80). [En ligne] <http://www.archive.org/details/graphicpresentat00brinrich> [lien vérifié le 02/08/2015].
- Brun J. (1993), La mobilité résidentielle et les sciences sociales. Transfert de concept et questions de méthodes, *Les annales de la recherche urbaine*, n°59-60, pp. 3-14.
- Brun P. (2003), Le récit de vie dans les sciences sociales. in : *L'écriture de la vie*, Dossier n°188. [En ligne] <http://www.revue-quartmonde.org/spip.php?article62> [lien vérifié le 02/08/2015]
- Brundson C., Charlton M. (2006), Local trend statistics for directional data – A moving window approach, *Computer, Environment and Urban Systems*, vol.30, pp. 130-142.
- Brunet, R. (1987), *La carte mode d'emploi*, Fayard-Reclus, 269 p.
-

-
- Brunet, R. (2009), Les sens de la distance, in : *La distance, objet géographique*, ATALA n°12, 32 p.
- Brunet, R., Ferras, R. et Théry, H. (1992), *Les mots de la géographie. Dictionnaire critique*, Reclus, La Documentation Française, Montpellier-Paris, 470 p.
- Bunge, W. (1962, 1966), *Theoretical geography*, (Second edition), Lund Studies in Geography, C : General and Mathematical Geography, n°1, Lund, Sweden, Gleerup.
- Bunge, W. (1969), Simplicity, in : *Geographical analysis*, vol. 1, n°3, pp. 88-91.
- Caquart, S. (2001), *Des cartes multimédias dans le débat public : pour une nouvelle conception de la cartographie appliquée à la gestion de l'eau*, Thèse de doctorat, Université Jean Monnet de Saint-Etienne, 277 p.
- Cattan, N., Grasland C. (1994), Migrations et effets de barrière en Tchécoslovaquie, in : Rey V. (ed.), *La Tchécoslovaquie en 1992 : transition, fragmentation, recomposition*, Presses de l'Ecole Normale Supérieure ENS, Fontenay-Saint-Cloud, Coll. Sociétés, Espaces, Temps, pp. 97-120.
- Cattan, N. (2004), Le monde au prisme des réseaux aériens, *Flux*, vol. 4, n°58, pp. 32-43. [En ligne] <http://www.cairn.info/revue-flux-2004-4-page-32.htm> [lien vérifié le 02/08/2015]
- Cauvin, C., Reymond, H, Serradj, A. (1978), *Discrétisation et représentation cartographique*, Reclus Mode d'emploi, 116 p.
- Cauvin C., Reymond H. (1984), Interaction spatiale et cartographie : les solutions de W. Tobler. *Espace, Populations, Sociétés*, n°3, 199, pp. 467-485. [En ligne] http://www.persee.fr/web/revues/home/prescript/article/espos_0755-7809_1991_num_9_3_1489 [lien vérifié le 02/08/2015]
- Cauvin C. (1998), Des transformations cartographiques, *M@ppemonde*, vol. n°49, n°98, 4 p. [En ligne] <http://www.mgm.fr/PUB/Mappemonde/M198/Cauvin.pdf> [lien vérifié le 02/08/2015]
- Cauvin, C., Enaud, C. (1991), De la cartographie de report à la cartographie transformationnelle. Proposition pour la représentation des données de population, *Espace, Population, Sociétés*, n°3, pp. 487-503.
- Cauvin, C., Escobar F., Serradj A. (2008a), *Cartographie thématique - Tome 1, Une nouvelle démarche*, Hermès Sciences Publications, Coll. Information Géographique, Traité IGAT, 248 p.
- Cauvin, C., Escobar F., Serradj A. (2008b), *Cartographie thématique - Tome 3, Méthodes quantitatives et transformations attributaires*, Hermès Sciences Publications, Coll. Information Géographique, Traité IGAT, Aspects fondamentaux de l'analyse spatiale, 282 p
- Certu (2008), Les recueils de données sur la mobilité urbaine, in *Note de synthèse : Mobilités et transports*, Hors Coll., n°2, 8 p.
- Chadule (groupe), (1987), *Initiation aux pratiques statistiques en géographie*, Masson, 4^{ème} éd., 203 p.
- Chapelon L. (1996), Modélisation multi-échelles des réseaux de transport : vers une plus grande précision de l'accessibilité. *M@ppemonde*, n°3, pp. 28-36. [En ligne] <http://www.mgm.fr/PUB/Mappemonde/M396/Chapelon.pdf> [lien vérifié le 02/08/2015]
- Chardonnel S. (2001), La time-geography : les individus dans le temps et dans l'espace, in : Sanders L., *Modèles en analyse spatiale*, Lavoisier, pp. 129-153.
- Chesneau, E. (2006), *Modèle d'amélioration automatique des contrastes de couleur en cartographie - Application aux cartes de risques*, Thèse de doctorat en Sciences de l'information géographique, Université Paris-Est Marne-la-Vallée France, 372 p.

-
- Cheyland, J.-P. (2007), Les processus spatio-temporels : quelques notions et concepts préalables à leur représentation, *M@ppemonde* vol. 87, n°3. [En ligne] <http://mappemonde.mgm.fr/num15/articles/art07303.html> [lien vérifié le 02/08/2015]
- Christophe S., Ruas A. (2009), Un processus de conception de légendes créatives sur mesure, in : Numéro spécial sur les publications de la conférence ICC'09, *Le monde des cartes*, Comité Français de Cartographie, n°202. [En ligne] <http://www.lecfc.fr/new/articles/202-article-6.pdf> [lien vérifié le 02/08/2015]
- Christophe, S. (2009), *Aide à la conception de légendes personnalisées et originales : proposition d'une méthode coopérative pour le choix des couleurs*, Thèse de doctorat en Sciences de l'information géographique, Université Paris-Est Marne-la-Vallée, 323 p. [En ligne] <http://tel.archives-ouvertes.fr/tel-00515333/fr/> [lien vérifié le 02/08/2015]
- Christophe, S., Zanin, C., Roussaffa H. (2011), Colours harmony in cartography, *Proceedings of the XXVe International Cartographic Conference (ICC'11)*, Paris, 3rd-8th July 2011.
- Chua, A., Marcheggiani, E., Servillo, L., Vande Moere, A. (2015), FlowSampler: Visual Analysis of Urban Flows in Geolocated Social Media Data, in: Aiello, L.M. and McFarland, D. (eds.), *SocInfo 2014 Workshops, Springer International Publishing Switzerland 2015*, LNCS 8852, pp. 5-7.
- Comité Français de cartographie (CFC) (1990), Glossaire de cartographie, in : *Le Monde des Cartes*, Revue du CFC, vol.123-124.
- Courgeau, D. (1973a), Migrants et migrations, *Population*, ed. de l'INED, n°28, pp. 95-129.
- Courgeau, D. (1973b), Migrations et découpage du territoire, *Populations*, vol. 28, n°3, pp. 511-537. [En ligne] http://www.persee.fr/web/revues/home/prescript/article/pop_0032-4663_1973_num_28_3_15462 [lien vérifié le 02/08/2015]
- Courgeau D. (1975a), Mesure et analyse des migrations, *Actes des Colloques Nationaux du CNRS, Analyse démographique et ses applications*, n°934, CNRS, pp. 151-161.
- Courgeau D. (1975b), Le concept de migration, *Actes du quatrième colloque de démographie africaine : Migrations – Etat Civil – Recensements Administratifs*, Ouagadougou, pp. 27-33.
- Courgeau D. (1980), *Analyse quantitative des migrations humaines*, Masson, Collection d'Anthropologie physique, Paris, 225 p.
- Courgeau D. (1988), *Méthodes de mesure de la mobilité spatiale, Migrations internes, mobilité temporaire, navettes*, ed. de l'INED, Paris, 301 p.
- Courgeau D. (2006), *Probabilités, démographie et sciences sociales, Mathématiques & Sciences humaines / Mathematics and Social Sciences*, n°167, pp. 27-50.
- Cruz, I. F., Tamassia, R. (n.d.), *Graph Drawing tutorial, Worcester Polytechnic Institute and Brown University*, 106 p. [En ligne] <http://www.cs.brown.edu/people/rt/papers/gd-tutorial/gd-constraints.pdf> [lien vérifié le 02/08/2015].
- (de) Saint Vaulry, A. (2008), *Base de données CHELEM – commerce international du CEPII, Rapport du CEPII*, n°8, 84 p. [En ligne] http://www.cepii.fr/PDF_PUB/wp/2008/wp2008-09.pdf [lien vérifié le 02/08/2015].
- Debarbieux, B. (2004), Cartes d'identités, cartes d'altérité, in : Bord J.-P., Baduel, R. (dir.), *Les cartes de la connaissance*, Karthala-Urbama, pp. 607-624.
- Denain, J.-C., Langlois, P. (1997), Anamorphose, analyse d'une métaphore spatiale, *Revue internationale de Géomatique (RIG)*, vol. 7, n°1, pp. 33-56.

- Denain, J.-C., Langlois, P. (1998), Cartographie en anamorphose, *M@ppemonde*, n°49, pp. 16-19. [En ligne] <http://www.mgm.fr/PUB/Mappemonde/M198/LangloisDenain.pdf> [lien vérifié le 2 août 2015].
- Dent, B. (1999), Dynamic representation: The design of flow map, Chapter 12, Part II: Techniques of quantitative thematic mapping, in: Dent, B., Cartography, *Thematic Map Design*, pp. 221-233. [Online] <http://gis.depaul.edu/shwang/teaching/geog360/private/FlowMap.pdf> [link checked, August 2, 2015]
- Dhée, F. (2013), *Amélioration de la carte topographique pour les daltoniens*, Thèse de doctorat en Géographie, Université Paris 1.
- Di Battista, G., Patrignani, M., Vargiu, F. (2000), A Split&Push Approach to 3D Orthogonal drawing, *Journal of Graph Algorithms and Applications*, vol. 4, n°3, pp. 105-133.
- Dorigo G., Tobler, W. (1983), Push Pull Migration Laws, *Annals of American Association of Geographers*, vol. 73, n°1, pp. 1-17.
- Dorling, D., F., L. (1991), *The Visualization of Spatial Social Structure*, Chapter 2 “People, Spaces and Places” and Chp. 6. “Cobweb of Flows”, Phd. Thesis, University of Newcastle upon Tyne, 577 p.
- Dorling, D. (2012)²²¹, *The Visualization of Spatial Social Structure*, Chichester, Wiley & Sons, Wiley Series, in: *Computational and quantitative social science*, vol. XLIX, 331 p.
- Drevelle, M. (2013), *D comme flux dominants*, Billet de blog du Groupe fmr. <http://groupefmr.hypotheses.org/1644>
- Drevelle, M. (2013), *D comme Degré*, Groupe fmr : <http://groupefmr.hypotheses.org/1425>
- Drevelle, M. (2013), *M comme flux majeur*, Billet de blog du Groupe fmr : <http://groupefmr.hypotheses.org/1406>
- Ducruet, C. (2010), *les mesures globales d'un réseau*. Méthodes, Synthèses du Groupe fmr. <https://halshs.archives-ouvertes.fr/halshs-00541902/document>
- Ducruet, C. (2011), *Simplification et partitionnement d'un graphe*, Méthodes, Synthèse du Groupe f.m.r., 8 p. [En ligne] https://halshs.archives-ouvertes.fr/file/index/docid/579065/filename/fmr6_partitionnement.pdf [lien vérifié le 02/08/2015].
- Dumolard, P. (1975), Région et régionalisation. [Une approche systémique] in : *l'Espace géographique*, Tome 4, n°2, pp. 83-111.
- Durand-Dastès, F. (2001), Les concepts de la modélisation en analyse spatiale, in : Sanders, L. (dir.), *Modèles en analyse spatiale*, Hermès Sciences, Lavoisier, Coll. Information Géographique et Aménagement du Territoire, pp. 31-59.
- Fairbairn, D. (2006), Measuring Map Complexity, *Cartographic Journal*, vol. 43, n°3, pp. 224 - 238.
- Fekete, J.-D. (2011), Visualizing Networks using Adjacency Matrices: Progresses and Challenges, Support of the presentation, *11th IEEE International Conference on Computer-Aided and Computer Graphics, CAD/Graphics '09*, Huangshan, pp. 27 p., 2011. [En ligne] <http://www.complexnetworks.fr/wp-content/uploads/2011/02/fekete-110224.pdf> [lien vérifié le 02/08/2015].
- Frémont, A., Frémont-Vanacore, A. (2014), Géographie des espaces maritimes, Dossier n°8104, *La documentation Photographique*, Dila, *La Documentation française*, pp. 18-31.

²²¹ L'ouvrage qui découle de cette thèse, publiée en 2012, est en partie disponible sur le site internet de l'auteur [Online] <http://www.dannydorling.org/books/visualisation/Homepage.html> [lien vérifié le 2 août 2015].

-
- Frémont, A., Frémont-Vanacore, A. (2015), Géopolitique des espaces maritimes, Dossier Mers et océans, *Carto, Le monde en cartes*, Areion Group, n° 28, pp. 13-23.
- Friedman, T. L. (2005), *The World Is Flat. A Brief History of the Twenty-first Century*, Farrar, Straus et Giroux, 421 p.
- Fruchterman, T., Reingold, E. (1991), Graph drawing by force-directed placement, *Software Practice Experience*, vol. 21, n°11, pp. 1129-1164.
- Fur (Le), A. (2000, 2007), *Pratiques de la cartographie*, Armand Colin, Coll. Synthèse, Série Géographie, n°97, 96 p.
- Fur (Le), A., Nonjon, A., Battistoni-Lemiere, A. (2010), *Cartes en mains. Méthodologie de la cartographie*, Ellipse Marketing, 247 p.
- Galton, A. P. (2003), Desiderata for a spatio-temporal geo-ontology, in: Kuhn, W., Worboys, M., F., Timpf, S. (Eds.), *Spatial information theory : Foundations of geographic information science*. Lecture Notes in Computer Science, vol. 2825, pp. 1-12.
- Ghonien, M., J., Fekete, J.-D., Castogliola, P. (2004), A comparison of the readability of graphs using node-link and matrix-based representation, *Proceedings of the IEEE Symposium on Information Visualization Infovis*, pp. 17-24.
- Ghonien, M., J., Fekete, J.-D., Castogliola, P. (2005), Readability of Graphs Using Node-Link and Matrix-Bases Representations: Controlled Experiment and Statistical Analysis, in : *Information Vizualisation Journal*, Palgrave Macmillan, vol. 2, n°4, pp. 114-135.
- Glennon, A., Goodchild, M. F. (2004)²²², *A GIS Flow Data Model*, White Paper, 3 p.
- Goodchild M. F., Yuan M., Covas J. T. (2007), Towards a general theory of geographic representation in GIS, *International Journal of Geographic Information Science*, vol. 21, n°3, pp. 239-260.
- Grasland, C. (1994), Limites politiques et barrières migratoires : l'exemple de l'ex-Tchécoslovaquie (1965-1989), in : Galluser W. (ed.), *Political Boundaries and Coexistence*, IGU Symposium, Basle, May 1994, pp. 425-440.
- Grasland, C., Bopda A. (1994), Migrations, régionalisations et régionalismes au Cameroun, *Espace, Populations, Sociétés*, n°1, pp. 109-129.
- Grasland, C. (1997), *Contribution à l'analyse géographique des maillages territoriaux*, Habilitation à Diriger des Recherches en géographie, Volume inédit, Université Paris I, 373 p.
- Grasland, C. (2004), *Interaction spatiale - Concepts*, Hypergéogé, <http://www.hypergeo.eu/spip.php?article2>
- Grasland, C., Baron, M. (2006), *Cartographie et Analyse spatiale*. Ecole d'été Statistiques, Yaoundé, août 2006, 46 p. [En ligne] <http://www.ums-riate.fr/ecoleyaounde2006/documents/fascicules/fluxcarto.pdf> [lien vérifié le 02/08/ 2015].
- Grasland, C., Didelon, C. (coord.) (2007), *Europe in the World*, ESPON project 3-4-1, Final Report, vol. 1, 243 p.
- Grasland, C., Bahoken, F., Beauguitte, L., Pion, G., Van Hamme, G. (2009), *Toolbox for flows and network analysis (Methodological Paper)*. Deliverable D.5.1. EuroBroadMap.Vision of Europe in the World. Small or medium scale focused project FP7-SSH-2007-1, 100 p.

²²² Le site internet dédié à cette version du logiciel *FlowMapper* de W. R. Tobler <http://dynamicgeography.ou.edu/flow/> développée par A. Glennon, n'est plus en activité [lien vérifié le 2 août 2015].

Grasland, C. (2009), Spatial Analysis of Social Facts, A tentative theoretical framework derived from Tobler's first law of geography and Blau's multilevel structural theory of society, 46 p. Publié in : Bavaud, F., Mager, C. (2010), *Advances in Quantitative and Theoretical Geography*, University of Lausanne. [En ligne] http://halshs.archives-ouvertes.fr/docs/00/41/06/69/PDF/Grasland_C_-_Spatial_Analysis_of_Social_Facts_-_23_Aug_2009.pdf [lien vérifié le 02/08/ 2015].

Grasland, C. (2010), *Transformation of flow matrix to symmetric Boolean graph: A proposal form comparative analysis*, EuroBroadMap.Vision of Europe in the World. Small or medium scale focused project FP7-SSH-2007-1, Flows and Network package, Methodological Paper n°3, 7 p.

Grasland, C., Van Hamme, G. (2010), La relocalisation des activités industrielles : une approche centre-périphérie des dynamiques mondiales et européennes, *l'Espace géographique*, Belin-Reclus, pp.1-19. [En ligne] <http://www.cairn.info/revue-espace-geographique-2010-1-page-1.htm> [lien vérifié le 11/10/ 2015].

Gataloup, C., Durant-Dastès, F. (1992), Le rôle des flux dans l'organisation des ensembles spatiaux, *L'information Géographique*, n° 56, pp. 35-52.

Gataloup, C., (2011), *Représenter le monde*, La Documentation Française, La Documentation photographique – Les dossiers, n° 8084, 64 p.

Gataloup, C., Fumey, G. (2015), *L'Atlas global d'un Monde global*, 150 p.

Guo, D. (2007), Visual analytics of spatial interaction patterns for pandemic decision, *International Journal of Geographic Information Science*, vol. 21, n°8/2007, pp. 857-877.

Guo, D. (2009), Flow mapping and multivariate visualization of large spatial interaction data, *Proceedings of the IEEE Symposium on Visualisation and Computer Graphics*, vol. 15, n°6, pp. 1041-1048.

Guo, D., Zhu, X. (2014), Origin-Destination flow data smoothing and mapping, *Proceedings of the IEEE Transactions on Visualization and Computer Graphics* [video] <https://vimeo.com/113238478>

Hägerstrand, T. (1957), Migrations and areas, in: Hanneberg D., Hägerstrand T., Odeving B. (eds), *Migrations in Sweden: a Symposium*, *Lund Studies in Geography*, n°B. 13, pp. 27-58.

Haggett, P. (1965²²³, 1973²²⁴), *Locational Analysis in Geography*, Arnold, Londres, 339 p. | *L'analyse spatiale en géographie humaine*, Armand Colin, 390 p.

Harness, H. D. (1838), *Atlas to Accompany the Second Report of the Railway Commissioners*, H.M.S.O, Dublin, Ireland

Harrower M. (2007), *Entry: Time, Time Geography, Temporal Change, and Cartography*, *The History of Cartography*, vol. 6 : The twentieth Century. Draft Paper, 6 p.

Henry, N., Fekete, J.-D. (2001), Matrix Explorer : un système pour l'analyse exploratoire des réseaux sociaux. [En ligne] http://research.microsoft.com/en-us/um/people/nath/docs/henry_ihm06.pdf [lien vérifié le 02/08/2015]

Henry, N., Fekete, J.-D. (2006), Matrix Explorer: a Dual-Representation System to Explorer Social Networks, *Proceedings of the IEEE Transactions on visualization and computer graphics*, vol. 12, n°5, 8 p.

²²³ La version anglaise de cet ouvrage a été rééditée en 1977 en deux volumes : *Locational Models* (vol. 1) et *Locational Methods* (vol. 2), 605 p.

²²⁴ Il s'agit de la date de publication de la version française de l'ouvrage.

-
- Henry, N., Fekete, J.-D., McGuffin, J. (2007), NodeTrix: A Hybrid Visualization of Social Networks, *Proceedings of the IEEE Transactions on Visualization and Computer Graphics*, vol. 13, n°6, pp. 1302-1039.
- Henry, N., Fekete, J.-D. (2009), Représentations visuelles alternatives pour les réseaux sociaux, *Réseaux*, n°152, pp. 59-92.
- Herman, I., Mélançon, G., Scott Marshall, S. (2000), Graph Visualization and Navigation in Information Visualization: A Survey, *Proceedings of the IEEE Transactions on Visualization and Computer Graphics*, vol. 6, n°1, pp. 24-43.
- Ho, Q., Nguyen H.-P., Jern, M. (2011), Implementation of a Flow Map Demonstrator for Analyzing Commuting and Migration Flow Statistics Data, *Procedia – Social and Behavioral Sciences, Spatial Thinking and Geographic Information Sciences*. [En ligne] <http://liu.diva-portal.org/smash/get/diva2:617988/FULLTEXT01.pdf> [lien vérifié le 04/08/2015]
- Holland, S., C., Plane, D., A. (2001), Methods of mapping migration flow patterns. *Southeastern Geographer*, vol. 41, n°1, pp. 89-104.
- Holten, D. (2006), Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data, *Proceedings of the IEEE Transactions on Visualization and Computer Graphics*, vol. 12, n°5, pp. 741-748.
- Holten, D., Isenberg, P., van Wijk, J., Fekete, J.-D. (2010), *Analysis, visualization, interaction*, Section Perception of Directed edges. Poster. [En ligne] <http://www.aviz.fr/wiki/uploads/Main/AVIZ-Poster-2012.png> [Lien vérifié le 02/08/ 2015].
- Holten, D., Van Wijks, J. (2009), Force-Directed Edge Bundling for Graph Visualization, in : Hege, H.-C., Hotz, I., Munzner, T., (Guest editor), *Journal compilation of The Eurographics Association /IEEE-VGTC Symposium on Visualization*, Vol.28, n°3, pp. 983-990.
- Inselberg, A. (1985), The Plane with Parallel Coordinates, *Visual Computer*, vol. 1, n°4, p. 69-91.
- Jegou, L., Deblonde, J.-P. (2012), Vers une visualisation de la complexité de l'image cartographique, *Cybergeo : European Journal of Geography, Rubrique Cartographie, Imagerie, SIG, document 600*, mis en ligne le 11 avril 2012. [En ligne] <http://cybergeo.revues.org/25271>[Lien vérifié le 02/08/ 2015].
- Jegou, L. (2013a), *Vers une nouvelle prise en compte de l'esthétique dans la composition de la carte thématique : propositions de méthodes et outils*, Thèse de doctorat en Géographie et Aménagement, Université Paul Sabatier, Toulouse, 462 p.
- Jegou L. (2013b), *Enquête cartographie et esthétique. Synthèse et commentaires des résultats – deuxième version*, Rapport de recherche non publié, 48 p.
- Joliveau, T. (2013), Chacun sa carte ? Le nouveau Google Maps, *M@ppemonde*, rubrique Internet. [En ligne] <http://mappemonde.mgm.fr/num38/internet/int13201.html> [Lien vérifié le 02/08/ 2015]
- Joly F. (1976), *La cartographie*, PUF, Coll. QSJ ?, 276 p.
- Kaddouri, L. (2008), Réflexion sur la sémiologie graphique animée des flux, *M@ppemonde*, n°89 (1-2008). [En ligne] <http://mappemonde.mgm.fr/num17/articles/art08104.html> [Lien vérifié le 02/08/ 2015]
- Kamada, T., Kawai, S. (1989), An algorithm for drawing general undirected graphs, *Journal Information Processing Letters*, vol. 31, n°. 1, pp. 7-15.
- L'Hostis, A. (1996), Transports et Aménagement du territoire: cartographie par images de synthèse d'une métrique réseau, *M@ppemonde*, n°3, pp. 37-43. [En ligne] <http://www.mgm.fr/PUB/Mappemonde/M396/LHostis.pdf> [lien vérifié le 2/08/2015]

-
- L'Hostis, A. (1997), Représentation en relief des réseaux de transport de l'arc atlantique : déformations de l'espace-temps en « vallées » et « montagnes », *Acte de colloque TheoQuant*, pp. 107-112. [En ligne] <http://thema.univ-fcomte.fr/theoq/pdf/1997/TQ1997%20ARTICLE%2012.pdf> [lien vérifié le 2/08/2015]
- L'Hostis, A. (2007), Graph theory and representation of distances: chronomaps, and other representations, in: Mathis, P., *Graphs and networks, multilevel modeling*, Lavoisier, London pp.177-191.
- L'Hostis A., Menerault P., Decoupigny C. (2004), Assessing spatial planning policy with accessibility indicators: the case of Lille's metropolis scenario, in: *Advances in spatial sciences, Transport developments and innovations in an evolving world*, Springer, Berlin, pp. 293-312.
- L'Hostis A., Leysens, T. (2012), *Les méthodes de mesure et de représentation de l'accessibilité dans les méthodes d'évaluation des projets de transport interurbains et périurbains de voyageurs: méthode, indicateurs, applications et limites de la contactabilité*, Rapport de recherche, 93 p. [En ligne] URL : <https://hal.archives-ouvertes.fr/hal-00734212/document> [Lien vérifié le 02/08/2015].
- Lambert, A., Bourqui, R., Auber, D. (2009), Winding Roads: Routing edges into bundles, *Proceedings of the Eurographics / IEEE-VGT Symposium on Visualization*, vol. 29, n°3, pp. 853-863.
- Lambert, A., Aubert, D., Mélançon, G. (2010), Living flows: enhanced exploration of edge-bundled graphs based on GPU-intensive edge rendering, *Proceeding of the 14th International Conference on Information Visualization*, pp. 523-530.
- Lanco-Bertrand S. (2005), *Analyse comparée des dynamiques spatiales des poissons et des pêcheurs : mouvements et distributions dans la pêcherie d'anchois (Engraulis Ringens) du Pérou*, Thèse de doctorat en Agronomie (halieutique), Ecole nationale supérieure d'agronomie de Rennes, IRD, 204 p.
- Langlois, P. (2003), Anamorphoses par transformations pseudo-équivalentes, *VIemes Rencontres de ThéoQuant*, 17 p.
- Lebart, L. (1969), Analyse Statistique de la Contiguïté, *Publications de l'ISUP*, n°XVIII, pp. 81-112.
- Mac Eachren, A. (1979), The evolution of thematic cartography / A research methodology and historical review, *The Canadian geographer*, vol. 16, n°1, pp. 17-33. [Online] http://www.geovista.psu.edu/publications/MacEachren/MacEachren_Evolution_1979.pdf [link checked, August 2, 2015]
- Mac Eachren, A. M. (1982), Map Complexity: Comparison and Measurement, *The American Cartographer*, vol. 9, n°1, pp. 31-46.
- Mac Eachren A. M. (1992), Visualizing uncertain information, *Cartographic Perspectives*, n°13, pp. 10-19.
- Mac Eachren, A. M., Ruxton, D., Taylor, F. (1994), *Visualization in modern cartography*, vol. 2, 345 p.
- Mac Eachren, A. M. (1995), *How Maps Work: Representation, Visualization and Design*, New York, The Guilford Press.
- Monmonier, M. (1993), *Comment faire mentir les cartes, Du mauvais usage de la géographie*, Flammarion, 233 p.
- Müller, J.-C. (1979), La cartographie d'une métrique non euclidienne : les distances-temps, *L'Espace géographique*, n°3, pp. 215-227.
- Müller, J.C. (1982), Non Euclidian geographic spaces: mapping functional distances, *Geographical Analysis*, vol. 14, n°3, pp. 189-203.

-
- Müller, J.-P. (1991), le problème des visualisations des migrations inter-Etats aux États-Unis, *Espace, populations, sociétés*, vol. 9, n°3, pp. 459-465.
- Nyusten, J. D., Dacey, M. F. (1961), A graph theory interpretation of nodal regions, *Papers and proceeding of the Regional Science Association*, vol. 7, n°1, pp. 29-42. [Online] http://deepblue.lib.umich.edu/bitstream/handle/2027.42/45977/10110_2005_Article_BF01969070.pdf?sequence=1 [link checked, August 2, 2015]
- O'Brien, R. (1992), *Global Financial Integration - The End of Geography*, Council on Foreign Relations Press, New York, 126 p.
- Otjacques, B. (2004), *Représentation graphique des interactions se produisant au cours d'un projet collaboratif dans le domaine de l'architecture*, Mémoire de DEA simulation et modélisation des espaces bâtis, Université Henri Poincaré, Nancy, 118 p.
- Palsky, G. (1996), *Des chiffres et des cartes. Naissance et développement de la cartographie quantitative française au XIXème siècle*, Rapport du Comité des Travaux Historiques, Ministère de l'Enseignement supérieur et de la Recherche Paris, Mémoire de la section de géographie, n°19, 331 p.
- Palsky, G. (2004a), Le temps des cartographes, in : Bord J.-P., Baduel, R., *Les cartes de la connaissance*, Karthala-Urbama, pp. 345-352.
- Palsky, G. (2004b), Le calcul par l'oeil, in : Bord J.-P., Baduel, R., *Les cartes de la connaissance*, Karthala-Urbama, pp. 588-605.
- Palsky, G. (2004c), *Carte choroplèthe*. Hypergéométrie. <http://www.hypergeo.eu/spip.php?article274>
- Palsky, G., Robic, M.-C. (2000, 2009), Aux sources de la sémiologie graphique, *Cybergeographie : European Journal of Geography, Dossier du Colloque 30 years of graphic semiology en l'honneur de Jacques Bertin*, Article 147 [En ligne] <http://cybergeographie.revues.org/554> [lien vérifié le 02/08/2015]
- Parks, M., J. (1987), *American Flow Mapping : A survey of the Flow Maps Found in the Twentieth Century Geography Textbooks, Including a Classification of the Various Flow Map Designs*, Unpublished Master Thesis, Georgia State University, Atlanta University.
- Passeron, J.-C. (1991), *Le raisonnement sociologique, L'espace non-poppérien du raisonnement naturel*, Nathan, pp. 111-133.
- Perin, C., Dragicevic, P. Fekete, J.-D. (2014). Revisiting Bertin's Matrices: New Interactions for Crafting Tabular Visualizations. *Proceedings of the TVCG, VIS' 2014*.
- Phan, D., Xiao, L., Yeh, R., Hanrahan, P., Winograd, T. (2005), Flow Map Layout, *Proceedings of IEEE Symposium on Information Visualization, INFOVIS 2005*, pp. 219-224.
- Poitevin, D. (1999), *La carte, moyen d'action. Guide pratique pour la conception et la réalisation de cartes*, Ellipses, 200 p.
- Poulain, M. (1991), Impact du nombre et de la taille des zones lorsqu'on estime un modèle spatial gravitationnel de migration, in : Pumain D. (ed), *Spatial analysis and population dynamics*, John Libbey Eurotext et INED, pp. 107-117.
- Poulain, M. (1981), *Contribution à l'analyse spatiale d'une matrice de migrations internes*, Thèse de doctorat, Université de Louvain La Neuve, Belgique, Cavey, 226 p.
- Pumain, D. (1986), Les migrations interrégionales de 1954 à 1982 : directions préférentielles et effets de barrière, *Population*, vol 41, pp. 378-389.

-
- Pumain, D., Saint-Julien, T. (1990), Les migrations interurbaines, *M@ppemonde*, n°90, n°1, pp. 33-34.
- Pumain, D., Saint-Julien, T. (2001), *Les interactions spatiales*, Armand Colin, Coll. Cursus Géographie, 191 p.
- Pumain, D. (2004) *Interaction*. Hypergéô. <http://www.hypergeo.eu/spip.php?article73>
- Rae, A. (2009), From spatial interaction data to spatial interaction information? Geovisualisation and spatial structures of migration from the 2001 UK census, *Computers, Environment and Urban Systems*, vol. 33, n°3, pp. 161-178. [Online] <https://mediamapping.wikischolars.columbia.edu/file/view/Rae+-+2009+-+From+spatial+interaction+data+to+spatial+interacti.pdf> [link checked, August 2, 2015]
- Raisz, E., J. (1938, 1948), *General cartography, McGraw-Hill Series in Geography*, New York, London.
- Raisz, E. (1944), Introduction, in : *Atlas of Global Geography*.
- Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Mauro, M., Claxton, R., Strogatz, H. (2010), Redrawing the map of Great Britain from a Network of Human Interactions. PLoS ONE, vol. 5, n°12. [Online] <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0014248> [link checked, August 2, 2015]
- Ravenstein E., G. (1876), The Birthplace of the People and the Laws of Migration, *The Geographical Magazine*, n°46, pp. 173-177, pp. 201-206, pp. 229-233.
- Ravenstein E., G. (1885), The Laws of Migration, *Journal of the Statistical Society of London*, n°52, pp. 167-235.
- Ravenstein E., G. (1889), The Laws of Migration, Second Paper, *Journal of the Statistical Society of London*, vol. 52, n°2, pp. 241-305.
- Renner, G., T. (1942), Human geography in the air age (a text for high school students), *The Macmillan company, Air Age Education Series (from the Aviation Education Research Group)*, Columbia University, New York, 258 p. [Online] <http://archive.org/details/humangeographyin00rennrch> [link checked, August 2, 2015]
- Richard ,Y., Zanin (Tobelem), C. (2009), L'Europe dans la régionalisation de l'espace mondial. Étude des flux commerciaux par un modèle d'interaction spatiale, *Géocarrefour, Quelle Europe vingt ans après la chute du Mur ?* Vol. 84/3, pp. 137-149. [En ligne] <http://geocarrefour.revues.org/7383?lang=en> [lien vérifié le 02/08/2015].
- Rimbert, S. (1990), *Carto-graphies*, Hermès, 176 p.
- Robinson, A., H. (1955), The 1837 map of Henry Drury Harness, *Geographical Journal*, n°121, pp. 440-450.
- Robinson, A., H. (1967), The Thematic maps of Charles Joseph Minard, *Imago Mundi* 1, pp. 95-108.
- Robinson, A., H. (1970), Erwin Josephus Raisz, 1893-1968, *Annals of the Association of American Geographers*, Vol. 60, n°1, pp. 189-193.
- Robinson, A., H. (1982), *Early thematic mapping in the history of cartography*, Chicago University Press, 266 p.
- Rosental, P.-A. (1997), Les formalisations spatiales de la mobilité. Fragments pour l'histoire longue d'une non réception, *Genèses*, n°29, pp. 75-98

Roth R., E., Wooddruff, A., W., Johnson, Z., F. (2010), Value-by-alpha maps: An alternative technique to the cartogram, *Cartographic Journal*, vol. 47, n°2, pp. 130-140.

Royer J-F. (2009), Migrants-migrations : un vieux thème revisité, *Actes des Journées de Méthodologie Statistique (JMS) de l'INSEE*.

Sanders, L. (1992), *Système de villes et synergétique*, Anthropos Economica, Coll. Villes, 274 p.

Schulten, S. (2012), Mapping the Nation, History and Cartography in Nineteenth-Century America, *University of Chicago Press*, 272 p.

Tabart, G., Athènes, S., Conversy, S., Vinont J.-L. (2007), Effets des paramètres graphiques sur la perception visuelle : expérimentations sur la forme, la surface, l'orientation des objets et la définition des écrans, *Actes de la Conférence Francophone sur l'Interaction Homme-Machine (IHM'2007)*, Paris, France, pp. 23-30. [En ligne] <https://hal-enac.archives-ouvertes.fr/hal-01021970/document> [lien vérifié le 02/08/2015].

Thomas, D., S. (1941), *Social and Economic Aspects of Swedish Population Movements: 1750-1933*, Macmillan Reference USA, New-York,

Tobler, W. R. (1967a), Of maps and matrices, *Journal of Regional Sciences* (supplement), n°72, pp. 275-280.

Tobler, W. R. (1967b), Computer use in Geography, *Behavioral Sciences*, n°12, pp. 57-58.

Tobler, W. R. (1969), Geographical Filters and their Inverses, *Geographical Analysis*, vol. 1, n°3, pp. 234-253.

Tobler, W. R. (1970a), A computer movie simulating urban growth in the Detroit region, *Economic Geography*, vol. 46, n°2, pp. 234-240.

Tobler, W. R. (1970b), Experiments in migration mapping by computer, *The American Cartographer*, vol. 14, n°2, pp. 155-163.

Tobler, W. R. (1975), Commodity Fields, *Proceedings of the IIASA conference*, 15 p.

Tobler, W. R. (1978), Migration Fields, in : Clark, W., Moore, E., (eds.), *Population mobility and residential change*, Coll. Studies in Geography, The William and Marion Haas Research Fund, Northwestern University, n°25, Evanston, pp. 215-232.

Tobler, W. R. (1979a), A geographical flow mapping program, *Geographical Analysis*, n°13, Vol. 1, pp. 1-20.

Tobler, W. R. (1979b), Smooth Pycnophylactic Interpolation for Geographical Regions, *Journal of the American Statistical Association*, vol.74, n°367, pp. 519-530.

Tobler, W. R. (1981a), Depicting federal fiscal transfer transferts, *The Professional Geographer*, vol. 33, n°4, pp. 419-422.

Tobler, W. R. (1981b), A model of geographic movement, *Geographical Analysis*, vol. 13, n°1, pp. 1-20.

Tobler, W. R. (1982), Cartographic study of movement tables, Presentation in the *National Computer Graphic Assembly - session on Statistical Graphics - Mapping*, Anaheim, 17-06-1982.

Tobler, W. R. (1983), Push-pull migration laws, *Annals of the American Association of Geographer*, n°73, pp. 1-17.

Tobler, W. R. (1987), Experiments in migration mapping by computer, *American Cartographer*, n°14, pp. 155-163.

Tobler, W. R. (1988), Resolution, resampling and all that, *Paper for Discussion at the International Geographical Union Global Database Planning Project*, Tynley Hall, England, May, 9-13, 1988, 9 p.

Tobler, W. R. (1990), Frame independent spatial analysis, in: Goodchild, M., Gopal, S. (eds), *Accuracy of Spatial Databases*, Taylor & Francis Group, London, pp. 115-122. [Online] http://www.geog.ucsb.edu/~tobler/publications/pdf_docs/Frame-independent-spatial-analysis.pdf [link checked, August 2, 2015]

Tobler, W. R. (1994, 1995), Migration Ravenstein, Thorntwaite, and Beyond. *Paper presented at the 1994 San Francisco' meeting of the Association of American Geographers (AAG'1994)* [Online] http://www.geog.ucsb.edu/~tobler/publications/pdf_docs/Migration-Ravenstein-Thornthwaite-and-beyond.pdf [link checked, August 2, 2015] | Published in 1995, in: *Urban Geography, Special Issue: Spatial Analysis in Geography*, vol. 16, n°4, pp. 327-343.

Tobler, W. R. (1997), Visualizing the impact of transportation on spatial relations, *Proceedings of the Hawaii's Western Regional Science Association meeting*.

Tobler, W. R. (1999), The World is Shriveling as it Shrinks. *Communication in an ESRI Users' Conference*, San Diego.

Tobler, W. R. (2000, 2002), Ma Vie: Grandir en Amérique et en Europe”, in : Gould, P., Bailly, A. (eds) *Mémoires de Géographes*, Anthropos, Paris, pp. 209–242. | Published in 2002 as : Tobler, W. R., 2002, Ma vie: Growing Up in America and Europe, in: Good, P., Pitts, L. R., (eds), *Geographical Voices: Fourteen Autobiographical Essays*, Syracuse University Press, Syracuse, pp. 293-322.

Tobler, W. R. (n.-d.1)²²⁵, *On viewing flow map*, Note de recherche non datée, 6 p.

Tobler, W. R. (n.-d. 2), *Les possibilités de la cartographie des mouvements*, Note de recherche non datée, 2 p.

Tobler, W. R. (n.-d. 3), *Optimal parsing of large arrays*, Note de recherche non datée, 3 p.

Tobler, W. R. (n-d), Pseudo-cartograms, *American Cartographer*, n°13, vol. 1, pp. 43-50.

Torricelli, G.-P. (1997), Navetteurs en Suisse et en Italie du Nord, *M@ppemonde*, n°4, pp. 9-13. [En ligne] <http://www.mgm.fr/PUB/Mappemonde/M497/Torricelli.pdf> [lien vérifié le 02/08/2015]

Tufte, E., R. (1986, 2001), *The Visual Display of Quantitative Information*, Graphics Press, Chelshire, Connecticut, 197 p.

Tufte, E., R. (1990), *Envisioning information*, Graphics Press, 127 p. [En ligne] <http://e-x-a.org/stuff/books/other/tufte,%20edward%20-%20envisioning%20information.pdf> [lien vérifié le 13/10/2015]

Tukey J. W. (1977), *L'analyse des données exploratoires*, Addison-Wesley, Reading, Massachusetts, 688 p.

Van Hamme, G., Grasland C., (coord.) (2011), *Divisions of the world according to flows and networks*, Work Package 5 “Flows and Networks” Synthesis from deliverable 5.8., 67 p.

²²⁵ Toutes ces notes de recherches de W. R. Tobler non datées sont disponibles dans les recueils de textes réunis par C. Cauvin, C. Grasland, avec la collaboration de R. Schlumberger (Publications 1959-2000) et sinon, sur le site de l'auteur.

-
- Verbeek, K., Buchin, K., Speckmann, B. (2011), Flow Map Layout via Spiral Trees, *Proceedings of the IEEE Transactions on Visualization and Computer Graphics*, vol. 17, n°12, pp. 2536-2544.
- Weger G. (1999), *Cartographie. Sémiologie graphique et conception cartographique*. vol. 1, Institut Géographique National, Ecole Nationale des Sciences Géographiques, 141 p.
- Wilkinson, L., Friendly, M., (2009), The history of the cluster heat map, *The American Statistician*, vol. 53, n°2, pp. 179-184.
- Wood, J., Dykes, A. (2010), Visualization of Origins, Destinations and Flows with OD Maps, *Cartographic Journal*, vol. 47, n°2, pp. 117-129.
- Wood, J., Slingsby, A. Dykes, A. (2011), Visualizing the Dynamics of London's Bicycle-Hire Scheme, *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 46, n°4, pp. 239-251.
- Xiao, N., Chun, Y.(2009), Visualizing migration flows using kriskograms, *Cartography and Geographic Information Science*, vol. 36, n°2, pp. 183-191.
- Zanin, C., Trémélo, M.-H. (2003), *Savoir-faire une carte. Aide à la conception et à la réalisation d'une carte thématique univariée*, Belin, Coll. Géographie, 200 p.
- Zanin, C., Lambert, N. (2009), *Atlas interactif des Régions européennes (AIRES)*, UMS 2414 RIATE, <http://aire.ums-riate.fr/fr#content/index.fr.php> [lien vérifié le 02/08/2015]
- Zanin, C., Ben Rebah, M. (2011), Les variables visuelles de la cartographie dynamique : analyse, critique et propositions, *Xème Rencontres de Théoquant, Nouvelles approches en géographie théorique et quantitative*, Université de Franche-Comté et de Bourgogne, Besançon, 23-25 février 2011.

Documents cartographiques

Seuls les documents remarquables sont listés. Ils sont classés par ordre alphabétique du nom du premier auteur. Leur intitulé est porté en italique (et non celui de l'ouvrage qui les accueille).

Agence d'Urbanisme de la région du Havre et de l'estuaire de la Seine (AURH) (2014), *Les interactions dominantes entre géozones par mode de transport – Analyse modale des interactions*, in : Atlas analytique, le transport de marchandises en Europe du Nord-Ouest – dynamiques territoriales, infrastructures et échanges interrégionaux, 69 p.

Bartolomew, J.-L. (1958), *World air routes*, in: Time Atlas of the World, Planche n°7.

Bertin, J. (1973), *Flux commerciaux de blé dans la méditerranée du XVe siècle*, in : Les diagrammes, les réseaux et les cartes, pp. 344.

Brinton & Ass. (eds.) (1939), *International Trade in Wheat Average 1927-28 – 1930-31*, in : Graphic Presentations, Brinton, W. C. & Associates, New York, 505 p [En ligne] <http://www.archive.org/details/graphicpresentat00brinrich> [lien vérifié le 02/08/2015].

Brinton & Ass. (eds.) (1939), *Migration Into and from North Dakota for the Period from 1920 to 1930*, in : Graphic Presentations, Brinton, W. C. & Associates, New York, 505 p [En ligne] <http://www.archive.org/details/graphicpresentat00brinrich> [lien vérifié le 02/08/2015].

Brinton & Ass. (eds.) (1939), *Routes of travel During Employment of 100 migratory-Casual Workers in Agriculture in the United States in 1934*, in : Graphic Presentations, Brinton, W. C. & Associates, New York, 505 p [En ligne] <http://www.archive.org/details/graphicpresentat00brinrich> [lien vérifié le 02/08/2015].

Butler, P. (2010), *Visualizing Friendships*, in Facebook engineering blog [En ligne] <https://www.facebook.com/notes/facebook-engineering/visualizing-friendships/469716398919>

Halley, E. (1701), Map of the known magnitude variations of the world, in: Dead Reckonings, Lost Arts in the Mathematical Science, Part I. [Online] <http://myreckonings.com/wordpress/2009/04/18/magnetic-deviation-comprehension-compensation-and-computation-part-i/> [link checked, August 2, 2015]

Harness, H. D. (1838), *Atlas to Accompany the Second Report of the Railway Commissioners*, H.M.S.O, Dublin, Ireland.

L'Hostis, A. (2000), *Le relief d'espace-temps de la route et de l'autoroute*, figure 10-9, in : L'Hostis A., Menerault P., Decoupigny C. (2004), Assessing spatial planning policy with accessibility indicators: the case of Lille's metropolis scenario, in: Advances in spatial sciences, Transport developments and innovations in an evolving world, Springer, Berlin pp. 293-312.

Mariot, P. (n.-d.), *Loisirs en Slovaquie*, in : Brunet, R. (1987), La carte mode d'emploi, pp. 89.

Minard, C.-J. (1862), *Carte figurative et approximative représentant pour l'année 1858 les immigrants du globe réalisée en 1862*.

Minard, C.-J. (1869), *Carte figurative des pertes successives en hommes de l'armée française dans la campagne de Russie, 1812-1813*.

Playfair, H. W. (1786), *Representing, by Means of Stained Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure and Debts of England during the Whole of the Eighteenth Century*. in: The commercial and political Atlas.

Raisz, E., 1944, *The Air Age*, in: Atlas of Global Geography, pp. 22-23.

Ravenstein, E., G. (1885), *currents of migration*, in : Ravenstein, E-G., *The Laws of Migration*, *Journal of the Statistical Society of London*, n°52, pp. 167-235.

Ravenstein, E-G. (1885), *Currents of migrations*, Map n°5, in : Ravenstein, E-G., *The Laws of Migration*, *Journal of the Statistical Society of London*, n°52, pp. 167-235.

Rekacewicz, P. (2012), *L'oeuf asiatique*, in : Dossier Cartes en colère, Atlas 2012 - Mondes Emergents, le Monde diplomatique. [En ligne] <http://blog.mondediplo.net/2013-12-22-Cartes-en-colere> [lien vérifié le 02/08/2015]

United States Department of Agriculture/ Bureau of Agricultural Economics (1939), *International Trade in Wheat. Average 1927-1928 – 1930-31*, in: Graphic Presentations, Brinton, W. C. & Associates, New York, 505 p [En ligne] <http://www.archive.org/details/graphicpresentat00brinrich> [lien vérifié le 02/08/2015].

United States National Resources Board, State Planning (1935), *A. Migration Into and from North Dakota for the Period from 1920 to 1930*, in: Brinton, W.-P. (eds), Graphic Presentation.

WPA, Division of Social Research (1937), *Routes of Travel During Employment of 100 Migratory-Casual Workers in Agriculture in the United States in 1933 and 1934*, in : The migratory Casual Workers, embedded in: Brinton, W.-P. (eds), Graphic Presentation.

Références techniques

Cette liste récapitule les applications interactives et les logiciels permettant la création de cartes de flux, mentionnés dans le texte et/ou utilisés pour réaliser les cartes présentées. Le classement est réalisé par ordre alphabétique du nom de l'application. Cette liste ne prétend pas à l'exhaustivité.

Nom	Type	Auteur(s)	Référence bibliographique et/ou lien internet [liens vérifiés le 13/10/2015]
<i>Bertifier</i>	Application en ligne	C. Perin et al. (INRIA)	Perin, C., Dragicevic, P. Fekete, J-D. (2014). Revisiting Bertin's Matrices: New Interactions for Crafting Tabular Visualizations. <i>Proceedings of the TVCG, VIS' 2014</i> . http://www.aviz.fr/bertifier
<i>Cartography</i> (including flows)	Package R	T. Giraud, N. Lambert	Commented Scripts to Build Maps with Rcartography https://cran.r-project.org/web/packages/cartography/vignettes/cartography.html
<i>Commodity flow mapper</i>	Application en ligne		http://129.252.37.169:8400/flowvis/commodity/index.html
<i>Distributive flow lines</i>	Boite à outils ArcGis 10x.	B. Gerlt, ESRI lab	http://blogs.esri.com/esri/apl/2013/08/26/flow-map-version-2/
<i>Flow Map Demonstrator</i>	Application libre	Q. Ho et al.	Ho, Q., Nguyen H.-P., Jern, M. (2011), Implementation of a Flow Map Demonstrator for Analyzing Commuting and Migration Flow Statistics Data, <i>Procedia – Social and Behavioral Sciences, Spatial Thinking and Geographic Information Sciences</i> . http://liu.diva-portal.org/smash/get/diva2:617988/FULLTEXT01.pdf
<i>Flow Map Layout</i>	Application libre		Phan, D., Xiao L., Ron Y., Hanrahan P., Winograd T., 2005, Flow Map Layout, <i>Stanford Computer Graphics Lab's archives of technical publications</i> , 6 p. [En ligne] http://graphics.stanford.edu/papers/flow_map_layout/flow_map_layout.pdf et http://graphics.stanford.edu/papers/flow_map_layout/ [liens vérifiés le 16 juin 2010]
<i>Flow Mapping Modul</i>	Boite à outils ArcGis10x.	K. Kamyoun.	Kamyoun K., Sang-II, L., Jungyeop, S., Eunyoung, C. (2012), Developing a Flow Mapping Module in a GIS Environment, <i>Cartographic Journal</i> , vol. 49, n°2, pp. 164-175.
<i>FlowMap</i>	Code JAVA		Code java pour création de cartes de flux dites

			« radiales » https://github.com/yifantastic/FlowMap
<i>Flowmap</i> (<i>huhugravity</i>)	Code d3		Code pour la réalisation de cartes de flux avec Leaflet et d3 https://github.com/huhugravity/flowmap
<i>Flowmap generator</i>	Boite à outils ArcGis10x.	S. Prasa	http://blogs.esri.com/esri/apl/2012/09/12/generating-distributive-flow-maps-with-arcgis/
<i>Flowmapper</i>	Plugin QGIS	C. Gulluoglu	Description : https://plugins.qgis.org/plugins/FlowMapper/ http://gis.stackexchange.com/questions/56666/how-does-flow-mapper-plugin-work Téléchargement : http://95.9.195.180/
<i>Flowmap software</i>	Logiciel propriétaire		Flowmap, développé par l'Université d'Utrecht http://flowmap.geo.uu.nl/
<i>FlowSampler</i>	Logiciel propriétaire		Chua, A., Marcheggiani, E., Servillo, L., Vande Moere, A. (2015), FlowSampler: Visual Analysis of Urban Flows in Geolocated Social Media Data, in Aiello, L.M. and McFarland, D. (eds.): <i>SocInfo 2014 Workshops</i> , Springer International Publishing Switzerland 2015, LNCS 8852, pp. 5-7.
<i>Flowstrates</i>	Logiciel libre	I. Boyandin	http://code.google.com/p/jflowmap/ https://diuf.unifr.ch/main/diva/sites/diuf.unifr.ch.main.diva/files/jflowmap-geova10.pdf
<i>Global migration - R</i>	Code R		Visualizing migration flows data with circular plots https://github.com/null2/globalmigration
<i>Integrate environment for spatial interaction data</i>	Application propriétaire	G. Danhuai	Danhuai, G. (2011), An Integrated Comparative Visual Analytics Environment for Spatial Interaction Data, Proceedings of the VIVA ACCESS'2011, Xi'an, China, 38 p. http://cybergis.cn/cn/fmapp/fm.html
<i>JFlowmap</i>	Logiciel libre	I. Boyandin	Visualisation de matrices de flux http://code.google.com/p/jflowmap/
<i>Mapping flows</i>	Code R	J. Cheshire	http://spatial.ly/2015/03/mapping-flows/ https://procomun.wordpress.com/2015/04/14/mapping-flows-in-r-with-data-table-and-lattice/

<i>Mapshaper</i>	Application en ligne	M. Bloch	Application permettant de généraliser des fonds de carte (shapefiles) http://www.mapshaper.org/
<i>Matrix Explorer</i>	Application propriétaire		Henry, N., Fekete, J.-D. (2006), MatrixExplorer: a Dual-Representation System to Explore Social Networks, <i>Proceedings of the IEEE Transactions on visualization and computer graphics</i> , vol. 12, n°5. http://research.microsoft.com/en-us/um/people/nath/docs/Henry_infovis06.pdf
<i>ODmap</i>	Application propriétaire	J. Wood et al.	Wood, J., Dykes, A. (2010), Visualization of Origins, Destinations and Flows with OD Maps, <i>Cartographic Journal</i> , vol. 47, n°2, pp. 117-129. http://vimeo.com/45078794
<i>Oursins</i>	Plugin QGIS	L. Cacheux	Plugin pour QGIS, à télécharger ici : https://plugins.qgis.org/plugins/Oursins/ Une description en est proposée par J. Drewes : http://www.dgisenra.nl/tutorial/flow-mapping-in-qgis/
<i>Radial flow map</i>	Boîte à outils ArcGis 10x.	M. Akella	Creating radial flow maps with ArcGIS http://blogs.esri.com/esri/arcgis/2011/09/06/creating-radial-flow-maps-with-arcgis/
<i>Trajpop</i>	Application en ligne	R. Cura	Cura, R. (2013), TrajPop (1.0 web version) [Web application]. http://trajpop.parisgeo.cnrs.fr/
<i>Visual tool</i>	Application propriétaire	G. et N. Andrienko	Andrienko, G., Andrienko, N., Wrobel, S., 2007, Visual Analytics Tools for Analysis of Movement Data, <i>ACM SIGKDD Explorations</i> , vol. 9, n°2, pp. 38-46.

Références personnelles

Seules sont listées mes références réalisées en lien direct avec mon sujet de thèse, hors littérature grise, publiées (ou en cours d'évaluation) en octobre 2015. Elles sont classées par ordre chronologique croissant. Celles qui sont précédées d'un astérisque sont mentionnées dans le texte.

Prix et distinctions

Premier prix obtenu avec O. Bonin et S. Mermet en 2013, du Concours de géovisualisation et cartographies dynamiques, Session Géovisualisation des liens et des flux, du Festival International de Géographie de Saint-Dié-des-Vosges (FIG 2013), pour l'application de géovisualisation interactive des matrices de flux CARTOD/MAPOD.

Second prix obtenu avec I. Bouhet - Andre-Poyaud en 2005, du Concours de l'exposition scientifique du Festival International de Géographie de Saint-Dié-des-Vosges (FIG 2005) pour la méthodologie statistique et cartographique mise en œuvre dans les cartes de l'Atlas de la présence étrangère en France en 1999.

Publications dans des revues classées par l'HCERES

Dubucs H., Dureau F., Giroud M., Imbert C., André-Poyaud I., Bahoken F., Les circulants entre métropoles européennes à l'épreuve de leurs mobilités. Une lecture temporelle, spatiale et sociale de la pénibilité, *Articulo Journal of Urban Research*, Dossier « Tiring mobilities », 7 | 2011, Online since 19 September 2010, <http://articulo.revues.org/1810> [link checked, August 2, 2015]

*Bahoken, F. (2012), Atlas of Sciences. Vizualizing what we knows », *M@ppemonde*, Rubrique A Voir, à Lire, n°104. [En ligne] <http://mappemonde.mgm.fr/num32/librairie/lib11401.html> [Lien vérifié le 04/08/2015]

*Bonin, O., Bahoken F., Mermet S. (2013), CARTOD/MAPOD, une application de géovisualisation de la demande de transport locale, *M@ppemonde*, Rubrique Vu au Festival International de Géographie – Concours de géovisualisation : palmarès 2013, n°112. <http://mappemonde.mgm.fr/num40/fig13/fig13401.html> [Lien vérifié le 04/08/2015]

Bahoken, F., Drevelle, M. (2013), Les réseaux dans le temps et dans l'espace. Synthèse de la seconde journée du groupe fmr, *Netcom*, n°27, 3/4, pp. 411-426. <http://netcom.revues.org/1488> [Lien vérifié le 04/08/2015]

Bahoken, F. (2014), L'intérêt du raisonnement logique en cartographie de flux. L'exemple de migrations internes, *European Journal of GIS and Spatial Analysis / Revue Internationale de Géomatique (RIG)*, Lavoisier, vol. 24, n°2, pp. 231-250.

Bahoken, F., (*sous presse*), Eléments pour une représentation (carto)graphique des matrices de flux, *M@ppemonde*.

Bahoken, F., (*sous presse*), L'approche cartographique de la décomposition des matrices de flux, *M@ppemonde*.

Publications dans des revues à comité de lecture non classées

Bahoken, F. (2005), De la présence camerounaise en France à « l'option diaspora », in : *Enjeux*, Bulletin de géopolitique pour l'Afrique centrale, n°24, pp. 9-10.

Bahoken F. (2009), Les documents cartographiques en ligne sur les migrations internationales, in : Dubus, G., Documents sur Internet et migrations internationales, *e-migrinter*, n°3, pp. 37-52. [En ligne] http://www.mshs.univ-poitiers.fr/migrinter/e-migrinter/200903/emigrinter2009_03_37.pdf

Bahoken F. (2009), Atlas sur les migrations internationales, in : Dubus, G., Documents sur Internet et migrations internationales, *e-migrinter*, n°3, pp. 53-57. [En ligne] http://www.mshs.univ-poitiers.fr/migrinter/e-migrinter/200903/emigrinter2009_03_53.pdf [lien vérifié le 02/08/2015].

*Bahoken, F. (2013a), A propos de la première carte de flux qui figure des flèches (Ravenstein, 1885), *Confins, revue franco-brésilienne de géographie*, n°17. [En ligne] <http://confins.revues.org/8187> [lien vérifié le 02/08/2015].

*Bahoken, F., Olteanu-Raimond, A.-M. (2013b), La construction d'une matrice de flux à partir de traces de téléphones portables, *Revue Cartes et Géomatique* du Comité Français de Cartographie, n°217, pp. 73-86. [En ligne] <http://www.lecfc.fr/new/articles/217-article-7.pdf> [lien vérifié le 02/08/2015].

*Bahoken, F. (2014), Pour le renouvellement de la sémiologie de la carte de flux, *Revue Cartes et Géomatique* du Comité Français de Cartographie, n°222, pp. 109-121.

Chapitres d'ouvrages

Bahoken, F., Bouhet - Andre-Poyaud, I. (2007), Cartographier les migrations internationales, in : Audebert Cédric, Ma Mung Emmanuel (Eds.), *Les nouveaux territoires migratoires : entre logiques globales et dynamiques locales*, 4ème partie, Ed. HumanitarianNet, Université de Deusto, Bilbao, pp. 267-269.

Bahoken F. (2007), Emigrés d'Afrique centrale : qu'en est-il, qui sont-ils et que font-ils ? in : Pondi J-E. (dir.), *Immigration et diaspora. Un regard Africain*, Maisonneuve & Larose / Africaine d'édition, pp. 31-53.

Imbert, C. (dir.), Dubucs, A., Dureau, F, Giroud, M., avec les contributions de Alarcão V., André-Poyaud I., Bahoken F., Chabrol M., Chapon J., Le Roux G., Lévy J-P., Malheiros J. (2014), *D'une métropole à l'autre. Pratiques urbaines et circulations dans l'espace européen*, Armand Colin, Coll. Recherches, 487 p.

Cartes de déplacements, de flux ou de mouvements

Bahoken F., De Corbier, B., Dubus, G., Gonin, P., Jourdain, L.-A., Kotlok, N., Lucas, I., Masse, M., Penegos-Pila, E., Prat, L., Vignon, G. (2005), *Regards pluriels : 44 activités pédagogiques sur les préjugés, le racisme et les migrations*, Poitiers, Orcades, non paginé, ISBN : 2-903809-40-2.

Olivier Clochard (dir.) (2012, 2013), *Atlas des migrants en Europe. Géographie critique des politiques migratoires*, (2e édition), réalisation d'une sélection de cartes de flux pour l'ouvrage publié chez Armand-Colin, Hors Coll., Réseau Migreurop, 144 p.

*Frémont, A., Frémont-Vanacore, A. (2014), Géographie des espaces maritimes, Dossier n°8104, *La documentation Photographique*, Dila, La Documentation française, pp. 18-31.

Frémont, A., Frémont-Vanacore, A. (2015), Géopolitique des espaces maritimes, Dossier Mers et océans, *Carto, Le monde en cartes*, Areion Group, n° 28, pp. 13-23.

Bahoken F., Samé-Ekobo, M. (2015), La Bataille du Cameroun, août 1914-mars 1916, *Carto, Le monde en cartes*, Rubrique Les grandes batailles, n°31, Areion Group, pp. 72-73.

Communications orales avec actes dans un congrès international ou national

Grasland C, Bahoken F., Beauguitte L., Geoffrey P., Van Hamme G. (2009) Globalization and Regionalization of the World Economy (1995-2005) : a theoretical and methodological approach, *16th European Colloquium on Quantitative and Theoretical Geography (ECQTG'09)*, National University of Ireland, Maynooth / National Centre for Geocomputation, 4-8 septembre 2009, Maynooth, Irlande.

Bahoken, F. (2011), Evolution des limites de régions fonctionnelles 'Européennes' liées aux flux internationaux, *EuroBroadMap International Conference*, Session C: Frontières, réseaux et régionalisations floues du monde, Rouen, France, 1-3 Décembre 2011.

Bahoken, F. (2011), Comparison of functional regionalization's of the world: a methodological study based on Intramax procedure, *17th European Colloquium on Quantitative and Theoretical Geography (ECQTG'11)*, Special Session "Networking the World: Concepts, methods, data", 2nd-5th September, Athens, Greece.

*Bahoken, F. (2011), Cartographie des flux et effets frontières. Colloque international Les frontières mobiles, Session spéciale « Cartographie des transfrontières », *XIèmes rencontres du réseau BRIT*, Genève/Grenoble, 6-9 septembre 2011.

Bahoken, F. (2012), Contribution du raisonnement logique à la cartographie des flux, *Actes du Colloque International d'Analyse Spatiale et de Géomatique (SAGEO'2013)*, Liège, Belgique, 6-9 novembre 2012, pp.13-29. → Communication sélectionnée pour une publication dans la Revue Internationale de Géomatique (RIG).

Bahoken, F., Beauguitte, L., Lhomme S. (2013), Visualizing world flows: a challenge between efficacy, accuracy and aesthetics, *19th European Colloquium on Quantitative and Theoretical Geography (ECQTG'013)*, Special Session 'Networking the world II: data, Practice and Visualizations', Paris-Dourdan, juillet 2013.

Olteanu-Raimond, A.-M., Bahoken, F., Couronne T., Smoreda, Z. (2013), Proposition de matrices de flux temporelles issues de l'activité d'individus mobiles, *Colloque International d'Analyse Spatiale et de Géomatique (SAGEO'2013)*, Brest, septembre 2013.

*Bahoken, F., Oltéanu-Raimond, A. (2013), From individual paths to OD flows with mobile tracks as a marker of migrations, *International Cartographic Conference (ICC'2013)*, From Pole to Pole, Session S7-A : Geospatial Analytics, Dresdes, 2013, august, 25-30. → Communication sélectionnée et publiée en français dans la revue *Cartes et Géomatique* du Comité Français de Cartographie, n°217, pp. 73-86.

Bahoken F. (2014), Pour le renouvellement de la sémiologie de la carte de flux. *Journées Enseigner la Sémiologie*, Thème 3 : La sémiologie, usages et renouveau, Commissions Enseignement et Sémiologie du Comité Français de Cartographie (CFC), Montpellier, 22-23 mai 2014. → Communication sélectionnée et publiée en français dans la revue *Cartes et Géomatique*, n°222, pp. 109-121.

Bahoken F., Grasland, C., Zanin, C. (2015), Flow mapping vs movement mapping: a theoretical and cartographic demonstration of their significance at a world scale, *19th European Colloquium in Theoretical and Quantitative Geography (ECTQG 2015)*, Bari, Italie, septembre 2015.

Communications orales sans actes, dans un congrès international, national ou un séminaire

Bahoken F., Bouhet - André-Poyaud I. (2005), Méthodologies pour l'analyse statistique et cartographique des migrations : l'exemple des Marocains en France. *Session d'études doctorales sur les migrations*, Meknès, Maroc, 12-16 Septembre.

Bahoken F., Gauvin, A. (2006), A la recherche de représentations carto-graphiques des trajectoires résidentielles des migrants internationaux, *Colloque international Vingt ans de recherche sur les migrations internationales*, Atelier 9 : Cartographier les migrations internationales : quelles données pour quelles représentations ?, Maison des Sciences de l'Homme et de la Société (MSHS), Poitiers, 3-8 juillet.

*Bahoken F. (2007), Les mobilités forcées de population : une approche cartographique, *Forum des doctorants Les Risques*, Ecole Doctorale de Géographie de Paris (ED-434), Institut de Géographie, Paris, 23 mars.

André-Poyaud I., Bahoken F. (2010), Améliorer la compréhension des pratiques de circulation à l'échelle biographique : une contribution (carto)graphique. *Séminaire de restitution des résultats du Projet MEREV*, Université de Lisbonne, Faculté de Lettres, Centre d'Etudes Géographiques (CEG-IGOT), Lisbonne, 8 Juillet.

Bahoken, F. (2012), Cartographie des flux et production d'indicateurs sur l'intensité des échanges, *Colloque « Les Interfaces : Enjeux de Nature, de Sciences et de Société (LIENS)*, Maison des Sciences de l'Homme et de la Société, Aix-en-Provence, 15-16 mai.

Bahoken, F., Olteanu-Raimond, A.-M., (2013c) Le contrôle du « filtrage spatial » en cartographie des flux ou comment se départir des découpages administratifs. *Les Rencontres de Géomatique de SIG-la Lettre: Quand la géomatique repousse les frontières*, Session spéciale DECOUP – au-delà des découpages administratifs, Cité Descartes, Marne-la-Vallée, juin 2013.

Bonin O., Bahoken, F., Mermet S., (2013), « CARTOD-MAPOD – une application de visualisation de la demande de transport locale ». *Festival International de Géographie, Salon de la géomatique*, Mini conférence Visualisation des liens et des lieux. Saint-Dié-des-Vosges, France, 4-6 octobre..

Bonin, O., Bahoken, F., Mermet S. (2013), Cartographie des flux, géovisualisation des liens, *Les Cafés Cartographiques*, 2 décembre.

Bonin O., Bahoken, F., Mermet S., (2014) « Visualisation d'une matrice Origine-Destination entre communes : problèmes cartographiques et régionalisation des flux ». *13e séminaire du Grand Equipement Base de Données (GEBD) Belgrand*, Méthodes d'analyse des flux domicile-travail. Cité Descartes, Marne-la-Vallée, France, 20 mars.

Bahoken, F. (2014) Proposition de solutions pour l'analyse des matrices Origine-Destination. *13e séminaire du Grand Equipement Base de Données (GEBD) Belgrand*, Méthodes d'analyse des flux domicile-travail. Cité Descartes, Marne-la-Vallée, France, 20 mars.

Bonin, O., Bahoken, F., Mermet S. (2014), Cartographie des flux origine-destination à l'échelle communale à partir du recensement de la population, *Les Rencontres de statistique appliquée, Les déplacements dans l'espace : Données, mesures, représentations*, Institut National d'Etudes Démographiques (INED) Paris, 28 novembre.

Comptes-rendu critiques d'articles

*Gian Paolo Torricelli, Navetteurs en Suisse et en Italie du Nord : images des flux et de leurs changements, *M@ppemonde*, 1997-4, pp 9-13. *Atelier Méthodes croisées de l'observation des mobilités*, UR 013, IRD, Bondy, 13-14 juin 2006

*Jean-Claude Müller, Le problème des visualisations des migrations inter-états aux Etats-Unis, *Espace, Populations, Sociétés*, 1991-3, pp 459-465. *Atelier Méthodes croisées de l'observation des mobilités*, UR 013, IRD, Bondy, 13-14 juin 2006.

Communications par affiches, dans un congrès international ou national

*Bahoken, F., Bouhet-André – Poyaud, I. (2005), *Rendre visible la présence étrangère en France en 1999. L'exemple des Marocains en France en 1999*. Festival International de Géographie (FIG'2005). Saint-Dié-Des-Vosges.

André-Poyaud, I., Bahoken, F. (2006), *Vagues migratoires d'arrivée des étrangers présents en France en 1999*. Poster scientifique présenté au Festival International de Géographie (FIG'2006). Saint-Dié-Des-Vosges.

*Bahoken, F. (2007), *Les mobilités forcées de population : une approche cartographique*. Poster scientifique présenté au Forum "Risques", Ecole Doctorale de Géographie de Paris, Institut de Géographie, Paris. Mars 2007.

Bahoken, F., Bonin, O., Mermet, S. (2014), *Introducing MAPOD: A tool for the geoviz of flow matrices* : in Transport Research Arena (TRA2014) International Conference, Outreach market place, Paris-La Défense, 14-17 avril 2014.

Autres documents

Bahoken F., 2011, Représentation graphique des matrices : graphe et/ou carte de flux ? Synthèse du groupe fmr (flux, matrices et réseaux), 13 p. http://halshs.archives-ouvertes.fr/docs/00/64/17/33/PDF/fmr9_representations.pdf [lien vérifié le 02/08/2015]

*Bahoken F., Beauguitte, L., Lhomme, S., 2013, La visualisation des réseaux. Principes, enjeux et perspectives. Synthèse du groupe fmr (flux, matrices et réseaux), 16 p. [En ligne] <https://halshs.archives-ouvertes.fr/halshs-00839905> [lien vérifié le 02/08/2015].

LISTE DES TABLES

Table des figures

Introduction générale –

Figure I-1-1. Une carte de flux	13
---------------------------------------	----

Chapitre 1 –

Figure 1-1. La structure générale d'une matrice.....	29
Figure 1-2. Une matrice de flux	30
Figure 1-3. Une matrice de flux et ses marges	31
Figure 1-4. Deux matrices de flux (migratoires et commerciaux)	33
Figure 1-5. Définition du flux dans un aquarium d'Hägerstrand	35
Figure 1-6. Illustration de la mesure du flux	36
Figure 1-7. Le prisme spatio-temporel résidentiel de Marie	42
Figure 1-8. Le décompte des migrations de Marie, selon différents intervalles temporels.....	45
Figure 1-9. Le décompte des migrations de Marie selon le découpage territorial	47
Figure 1-10. Illustration du franchissement d'une limite de zone.....	49
Figure 1-11. La trajectoire spatiale de Marie dans le plan	49
Figure 1-12. L'effet du franchissement des limites de zones.....	50
Figure 1-13. Les matrices (régionales et communales) issues des migrations de Marie.....	52
Figure 1-14. Les matrices complémentaires.....	53

Chapitre 2 –

Figure 2-1. Le découpage en 114 districts de l'ex-Tchécoslovaquie de 1990.....	57
Figure 2-2. La matrice ordonnée de l'ex-Tchécoslovaquie (1990)	58
Figure 2-3. Le découpage régional sur la matrice ordonnée de l'ex-Tchécoslovaquie (1990)	59
Figure 2-4. La matrice en points proportionnels de l'ex-Tchécoslovaquie (1990)	62
Figure 2-5. La matrice en étoiles de l'ex-Tchécoslovaquie (1990).....	63
Figure 2-6. Exemples de matrices ordonnées hybrides	64
Figure 2-7. La matrice en 3D de l'ex-Tchécoslovaquie (1990)	65
Figure 2-8. Le graphe des relations observées de l'ex-Tchécoslovaquie (1990)	68
Figure 2-9. Le graphe des relations interdistricts de l'ex-Tchécoslovaquie (1990)	69
Figure 2-10. Le graphe des relations interdistricts de l'ex-Tchécoslovaquie (1990) – variante	70
Figure 2-11. Les différents types de graphes.....	71
Figure 2-12. Exemple de représentation des flux majeurs et dominants.....	72
Figure 2-13. Exemple de contraintes scalaires	72
Figure 2-14. Exemple d'application de forces directes moyennes et d'attracteurs	73

Figure 2-15. Exemples de graphes des relations interdistricts de l'ex-Tchécoslovaquie supérieures à 100, contraintes par la position des nœuds.....	74
Figure 2-16. Exemples de contraintes esthétiques de placement optimal des nœuds	75
Figure 2-17. Exemples de contraintes sémiologiques sur le type de liaisons.....	76
Figure 2-18. Illustration du diagramme <i>GeneaQuilts</i> (Fekete, 2010)	77
Figure 2-19. Les deux familles de graphes non planaires	78
Figure 2-20. Une carte des flux de l'ex-Tchécoslovaquie en 1990 (version 1).....	85
Figure 2-21. Présentation de <i>l'effet-spaghetti</i> caractéristique de la cartographie des matrices de flux denses	87
Figure 2-22. Cartographie des valeurs de flux supérieures à 100 migrants, dans l'ex-Tchécoslovaquie, de 1990	88
Figure 2-23. Cartographie des flux dominants des migrations de l'ex-Tchécoslovaquie, en 1990.....	89
Figure 2-24. Cartographie du volume bilatéral de migrations de l'ex-Tchécoslovaquie en 1990.....	92
Figure 2-25. Les procédés de représentation des flux, du point de vue des lieux	95
Figure 2-26. Une carte des flux de l'ex-Tchécoslovaquie de 1990, en symboles plus-moins.....	95
Figure 2-27. Une carte des migrations tchécoslovaques (1990) interpolée.....	96
Figure 2-28. La figure de la carte de type <i>Pycnophylactic</i> (Tobler, 1979).....	96
Figure 2-29. Une carte de l'attractivité des flux de l'ex-Tchécoslovaquie de 1990, en anamorphose scalaire.....	97
Figure 2-30. Illustration des flux sur un fond d'anamorphose scalaire (Dorling, 1991).....	99
Figure 2-31. Une double carte de flux combinée à une matrice ordonnée	101
Figure 2-32. La représentation de flux sous la forme de kriskogrammes (Xiao, et al., 2009)	102
Figure 2-33. Illustrations d'une représentation des flux sous forme d'arbre hiérarchique.....	103
Figure 2-34. Un graphe des flux au départ de Bratislava (ex- République Tchèque), en 1990.....	104
Figure 2-35. Cartographie du volume de flux de l'ex-Tchécoslovaquie (1990), selon le procédé de <i>l'edge bundling</i> hiérarchique.....	105
Chapitre 3 –	
Figure 3-1. La transformation de la matrice de flux (Fij) en matrice de liens (Lij)	113
Figure 3-2. Illustration de la cartographie de l'information disponible (Lij)	114
Figure 3-3. Cartographie de l'information disponible (Lij) dans la matrice de flux de l'ex-Tchécoslovaquie (1990)	114
Figure 3-4. Du tableau de contingence à la matrice de flux	116
Figure 3-5. Cartographie des probabilités d'émission (en %) de l'ex-Tchécoslovaquie (1990)	119
Figure 3-6. Illustration du caractère symétrique (et asymétrique) d'une matrice.....	121
Figure 3-7. Illustration d'une procédure de symétrisation d'une matrice (Fij)	122
Figure 3-8. Illustration de la cartographie d'une matrice de flux symétrique	122
Figure 3-9. Illustration de la procédure d' <i>antisymétrisation</i> de la matrice (Fij)	123
Figure 3-10. Illustration de la cartographie d'une matrice de flux antisymétrique	124
Figure 3-11. Illustration de la décomposition de la matrice (Fij) en composantes (Fij ⁺) et (Fij ⁻)	126

Figure 3-12. Illustration de la cartographie classique (liens, lieux) de la matrice (Fij).....	128
Figure 3-13. Cartographie des flux de l'ex-Tchécoslovaquie (1990).....	129
Figure 3-14. Cartographie de la composante symétrique (Fij ⁺) des flux (version 1)	130
Figure 3-15. Cartographie du <i>volume bilatéral</i> (Fij ⁺) de flux de l'ex-Tchécoslovaquie (1990).....	130
Figure 3-16. Illustration de la variation de la valeur du solde, en fonction de la formule.....	131
Figure 3-17. Illustration de la cartographie classique du transfert net (Fij)	132
Figure 3-18. Illustration de la cartographie du transfert net (Fij ⁻) (variante en deux cartes).....	132
Figure 3-19. Cartographie du <i>transfert net</i> (Fij ⁻) < 0 des flux de l'ex-Tchécoslovaquie (1990)	133
Figure 3-20. Proposition de cartographie du <i>transfert net</i> des flux de l'ex-Tchécoslovaquie (1990) : (Fij ⁻) > 0	134
Figure 3-21. Illustration de l'asymétrie d'une matrice de flux (Fij).....	136
Figure 3-22. Illustration de la cartographie classique de l'asymétrie des flux	137
Figure 3-23. Cartographie du <i>volume</i> et de l' <i>asymétrie</i> des flux de l'ex-Tchécoslovaquie (1990).....	137
Chapitre 4 –	
Figure 4-1. Le premier graphique statistique sur des flux (W. Playfair, 1786).....	144
Figure 4-2. La première carte de flux (H.D. Harness, 1836).....	145
Figure 4-3. La carte de C.-J. Minard (1869) sur la campagne napoléonienne de Russie	147
Figure 4-4. La gestion de la couleur, selon C.-J. Minard (1862).....	149
Figure 4-5. Illustration d'un diagramme Sankey (2015)	151
Figure 4-6. Exemple de carte réalisée selon le principe du diagramme Sankey	152
Figure 4-7. La première carte de flux où figurent des flèches (d'après Ravenstein, 1885).....	154
Figure 4-8. Illustration de l'anamorphose vectorielle multipolaire	157
Figure 4-9. L'ère aérienne et le rétrécissement du monde	159
Figure 4-10. Une représentation globale des flux et réseaux, selon E. Raisz (1944)	160
Figure 4-11. Les routes aériennes mondiales	161
Figure 4-12. Les champs migratoires	164
Figure 4-13. La direction préférentielle des flux	165
Figure 4-14. Deux figures de la carte de flux de la première moitié du XX ^e siècle	166
Figure 4-15. Quelques figures actuelles de la carte de flux (~2010).....	171
Figure 4-16. Une figure de la carte de flux (1939).....	171
Figure 4-17. La cartographie de routes, seconde moitié du XX ^{ème} siècle (Bertin, 1973).....	173
Figure 4-18. La carte des amitiés de Facebook©	174
Figure 4-19. Exemple de carte réalisée par E. Halley (1701).....	181
Figure 4-20. Illustration d'une série de variantes autour du trait droit, proposées par B. Dent (1999).....	182
Figure 4-21. Exemples de cartes de flux en trois dimensions, d'après O. Klein et C. Cauvin (2011)	183
Figure 4-22. Illustration d'une version du trait incurvé et en 3D	184
Figure 4-23. Illustration du principe du « ressort », par A. L'Hostis (2003).....	185
Figure 4-24. Exemple de “chronocartes” (L'Hostis, 2000, 2003).....	186

Figure 4-25. Les types de flèches	187
Figure 4-26. Illustration de l'effet de convergence (et de divergence) des flux	188
Figure 4-27. Illustration des problèmes de place des flèches et des solutions existantes.....	190
Figure 4-28. Illustration des possibilités de disposition des flèches.....	191
Figure 4-29. Illustration de la <i>défonce graphique</i> appliquée aux flèches.....	191
Figure 4-30. Illustration de la mise à l'échelle des figurés.....	193
Figure 4-31. Cartographie de flux selon le procédé du triangle (Bertin, 1973).....	194
Figure 4-32. Cartographie de flux issue du projet <i>Cartogram</i> (Torricelli, 1997).....	196
Figure 4-33. Illustration du procédé de l'oursin.....	197
Chapitre 5 –	
Figure 5-1. Illustration du caractère facultatif de la teinte, pour cartographier le volume de flux (Fij ⁺)	215
Figure 5-2. Vers une nouvelle esthétique de la carte de flux	216
Figure 5-3. 'L'œuf asiatique' (Rekacewicz, 2012).....	218
Figure 5-4. L'esthétique de la couleur noire	219
Figure 5-5. Sélection de cartes du réseau de métro parisien proposées sur un fond sombre, dans le cadre du concours de cartographie <i>Check my Map</i> ©	221
Figure 5-6. Illustration de la mobilisation du canal alpha de la teinte.....	225
Figure 5-7. Cartographie renouvelée de la matrice (Fij ⁺) de l'Isère	225
Figure 5-8. Expérimentations sur l'esthétique du noir, à partir de la carte du volume de flux	227
Figure 5-9. Les réponses aux questions F10 et C6 de l'enquête Cartographie et Esthétique (Jégou, 2013).....	228
Figure 5-10. Les principaux motifs du volume de flux de l'Isère en 2010	229
Figure 5-11. Illustration de la mise en œuvre, sur des flux, des nuances de teintes opposées	232
Figure 5-12. Le renouvellement de la sémiologie classique du solde bilatéral (Fij ⁻).....	233
Figure 5-13. Cartographie renouvelée de la matrice (Fij ⁻) de l'Isère	234
Figure 5-14. Le découpage du monde, selon CHELEM	236
Figure 5-15. Cartographie de liens commerciaux internationaux (Zanin, Richard, 2009).....	237
Figure 5-16. Représentation graphique des échanges de produits chimiques entre grandes régions mondiales (Grasland, Van Hamme, 2010)	238
Figure 5-17. Cartographie des liaisons commerciales (Lij) existantes.....	239
Figure 5-18. Cartographie des volumes d'échanges commerciaux (Fij ⁺)	240
Figure 5-19. Cartographie du solde bilatéral (Fij ⁻) des échanges commerciaux mondiaux	241
Figure 5-20. Cartographie des flux commerciaux mondiaux d'exportations (Fij) ou d'importation (Fji)	242
Figure 5-21. Cartographie du solde bilatéral des échanges commerciaux (Fij ⁻)	243
Chapitre 6 –	
Figure 6-1. Cadre sémiotique de réflexion sur la construction de la carte de flux : aspects cartographiques et sémantiques.....	249

Figure 6-2. Présentation des données de l'exemple des flux mondiaux.....	251
Figure 6-3. Cartographie des flux internationaux (volume bilatéral) de marchandises.....	252
Figure 6-4. Illustration du volume bilatéral (F_{ij}^+) de flux internationaux de marchandises.....	253
Figure 6-5. La construction d'une matrice d'intensité (I_{ij}) de flux	256
Figure 6-6. Exemples de solutions liées à la transformation de l'apparence des liens.....	259
Figure 6-7. L'effet du changement de l'apparence des liens et de la projection cartographique sur la perception des flux internationaux de marchandises.....	260
Figure 6-8. Illustration des distances géométriques : géographiques et euclidiennes (Googlemaps, 2015).....	264
Figure 6-9. Illustration de la variation de la perception des quantités échangées à l'échelle mondiale, l'exemple du trafic mondial de conteneurs (2012).....	270
Figure 6-10. La concordance visuelle des distances <i>cartographiques</i> , <i>géographiques</i> et <i>empiriques</i> perçues à l'échelle locale, l'exemple du volume bilatéral (F_{ij}^+) de navettes de l'Isère.....	271
Figure 6-11. La construction de la matrice de <i>mouvement spatial</i> $M_{ij}=(F_{ij}.D_{ij})$	273
Figure 6-12. L'analogie du <i>mouvement</i> avec une course à pied.....	274
Figure 6-13. Illustration de la cartographie d'une <i>logique de flux</i> versus <i>de mouvement</i>	276
Figure 6-14. Exemple de carte de flux de marchandises, exprimée en tonnes kilométriques	277
Figure 6-15. Cartographie de flux <i>versus</i> cartographie de mouvement.....	279
Figure 6-16. Cadre sémiotique de réflexion sur la construction de la carte de flux (complété).....	280
Chapitre 7 –	
Figure 7-1. Cartographie des liaisons (L_{ij}) de l'Isère.....	288
Figure 7-2. L'analogie avec la <i>densité de mouvement</i> des coureurs.....	290
Figure 7-3. Présentation d'une matrice de flux mondiaux de passagers (OACI, 2000).....	292
Figure 7-4. Illustration de la segmentation d'une matrice de flux, pour une collection de cartes.....	293
Figure 7-5. Collection de cartes du volume bilatéral de l'Isère (F_{ij}^+) issue d'une discrétisation de la matrice des distances (D_{ij}) selon les moyennes emboîtées	295
Figure 7-6. Collection de cartes du volume bilatéral de l'Isère (F_{ij}^+) issue d'une discrétisation de la matrice des distances (D_{ij}) selon les moyennes emboîtées	296
Figure 7-7. Collection de deux cartes du volume bilatéral des flux (F_{ij}^+) de l'Isère, issue d'une segmentation de la matrice des distances (D_{ij}) en fonction de la médiane	297
Figure 7-8. Collection de quatre cartes du volume bilatéral des flux (F_{ij}^+) de l'Isère, issue d'une segmentation de la matrice des distances (D_{ij}) selon les quartiles.....	298
Figure 7-9. Illustration de la cartographie de flux entre des zones limitrophes ($C_{ij}=1$).....	301
Figure 7-10. Graphe de voisinage ($C_{ij}=1$) des communes de l'Isère	301
Figure 7-11. Cartographie des flux (F_{ij}^+) \leftarrow ($C_{ij}=1$) entre les communes limitrophes de l'Isère.....	302
Figure 7-12. Présentation de la matrice exemple local à huit lieux.....	303
Figure 7-13. Illustration de la construction d'une matrice de contigüité ordinale (CK_{ij})	303
Figure 7-14. Graphes de voisinage des communes de l'Isère (CK_{ij}) avec $k=1$ et $k=2$	304
Figure 7-15. Cartographie des flux de l'Isère (F_{ij}) \leftarrow [CK_{ij} , $k=1$ et $k=2$]	305

Figure 7-16. Cartographie des flux de l'Isère (F_{ij}) \leftarrow [(CK $_{ij}$), $k=4$]	306
Figure 7-17. Illustration de la construction d'une matrice d'appartenance (A_{ij})	307
Figure 7-18. Cartographie des flux intercommunaux et intra bassins de vie de l'Isère	308
Figure 7-19. Cartographie des flux intercommunaux et inter bassins de vie de l'Isère	309
Figure 7-20. Volumes bilatéraux de flux inter bassins de vie de l'Isère	310
Figure 7-21. La vision polaire des flux aériens internationaux de passagers en 2000 (version 2)	312
Figure 7-22. La vision polaire des flux aériens internationaux de passagers en 2000 (version 3)	313
Figure 7-23. Cartographie du volume de flux de passagers internationaux en petits courriers.....	314
Figure 7-24. Cartographie du volume de flux de passagers internationaux entre pays voisins.....	316
Figure 7-25. Cartographie du volume de flux de passagers internationaux, sur l'espace euro méditerranéen, en petits courriers.....	317
Figure 7-26. Le découpage du monde en sept régions du niveau WUTS-2.....	319
Figure 7-27. Cartographie du volume de flux inter macro régionaux mondiaux de passagers aériens	321
Figure 7-28. Cartographie du volume de flux internationaux inter régionaux de passagers aériens...	322
Chapitre 8 –	
Figure 8-1. Expérimentation de la sélection des <i>Top-n</i> flux (M_{ij}^+) de l'Isère (2010)	327
Figure 8-2. L'histogramme des valeurs de flux (M_{ij}^+) de l'Isère.....	328
Figure 8-3. La courbe de distribution théorique des valeurs de flux	329
Figure 8-4. Histogrammes des valeurs transformées des matrices (F_{ij}^+) et (M_{ij}^+) de l'Isère.....	330
Figure 8-5. Diagramme Q-Q des valeurs de $\log(F_{ij}^+)$ et $\log(M_{ij}^+)$ de l'Isère	331
Figure 8-6. Les courbes de concentration de flux (F_{ij}^+) et de mouvements (M_{ij}^+) de l'Isère	332
Figure 8-7. Exploration « interactive » de l'effet cartographique de l'application du double critère de sélection sur la matrice (M_{ij}^+)	334
Figure 8-8. Cartographie des 80% de flux <i>significatifs</i> des matrices (F_{ij}^+) et (M_{ij}^+) de l'Isère	338
Figure 8-9. Cartographie des flux de l'Isère, selon différents objectifs de densité de la matrice.....	340
Figure 8-10. Architecture générale d'une application en ligne de sélection interactive de l'information de flux à cartographier.....	342
Chapitre 9 –	
Figure 9-1. La matrice de flux téléphoniques de (Nyusten et Dacey, 1961)	351
Figure 9-2. La matrice d'adjacence des flux téléphoniques de (Nyusten et Dacey, 1961)	352
Figure 9-3. Le graphe nodal des flux téléphoniques (Nyusten et Dacey, 1961)	353
Figure 9-4. Cartographie de la structure nodale des flux téléphoniques (Nyusten et Dacey, 1961) ...	354
Figure 9-5. Sémiologie de Nyusten-Dacey <i>versus</i> sémiologie usuelle.....	355
Figure 9-6. Les quatre formes corrigées de la sémiologie cartographique de Nyusten-Dacey	356
Figure 9-7. Illustration de la sélection de <i>flux majeurs</i> de (F_{ij}).....	358
Figure 9-8. Illustration de la cartographie des <i>flux majeurs</i> de (F_{ij})	359
Figure 9-9. Illustration de la cartographie de l'attractivité des lieux (D_j).....	361
Figure 9-10. Illustration de la cartographie des liens (L_{ij}) issus des flux dominants (F_{id}).....	361

Figure 9-11. Illustration de la hiérarchie des lieux du système de flux dominants (Fid)	363
Figure 9-12. Illustration de la régionalisation issue du système de flux (Fij) dominants (Fid).....	365
Figure 9-13. Illustration des sous-bassins régionaux issus du système de flux (Fij) dominants (Fid)	366
Figure 9-14. Graphe des relations nodales de la matrice (Fij) de l'Isère.....	368
Figure 9-15. Cartographie des pôles nodaux et des sous-bassins d'emploi du système de flux (Fij) dominants (Fid) de l'Isère	370
Figure 9-16. Régions et sous-régions du système de <i>flux dominants</i> (Fid) de l'Isère (Fij).....	371
Figure 9-17. Le découpage de l'Isère en Zone d'Emplois, selon l'INSEE (2011).....	372
Figure 9-18. Comparaison des principaux bassins régionaux issus des flux dominants (Fid) avec les Zones d'emplois de l'INSEE, pour l'Isère	373
Figure 9-19. Graphe des liaisons commerciales nodales issues de l'analyse des <i>flux dominants</i> (Fij) (méthode Nyusten-Dacey, 1961).....	376
Figure 9-20. Graphe des régions et sous-régions nodales issues de l'analyse des <i>flux dominants</i> du commerce mondial d'importation (Fij) de l'OMC (2007)	378
Figure 9-21. Arborescence principale des pays formant la région polarisée par l'Allemagne.....	379
Figure 9-22. Cartographie des deux principales régions mondiales issues des flux dominants (Fij) commerciaux internationaux	381
Figure 9-23. Cartographie du découpage en sous-régions de l'aire d'influence américaine issue des <i>flux dominants</i> (Fij) commerciaux internationaux.....	382
Figure 9-24. Cartographie de l'attractivité des lieux (Dj), par le recours à la <i>Taille</i>	383
Figure 9-25. Cartographie en anamorphose de la capacité d'importation (Dj) des pays.....	384
Figure 9-26. Cartographie des valeurs de <i>flux dominants</i> (Fij) du commerce international, sur un fond de carte anamorphosé	385
Figure 9-27. Illustration de la variation de la cartographie des <i>flux dominants</i> (Fij) versus (Fji)	387
Figure 9-28. Graphe des régions et sous-régions nodales issues de l'analyse des flux dominants du commerce mondial d'exportation (Fij) de l'OMC (2007).....	388
Figure 9-29. Illustration de la cartographie de <i>flux dominants</i> (Fij ⁺)	389

Table des tableaux**Chapitre 1 –**

Tableau 1-1. La variation du statut de migrante de Marie en fonction des réponses	43
---	----

Chapitre 3 –

Tableau 3-1. Indicateurs de fréquence simple des flux	118
---	-----

Tableau 3-2. Indicateurs de fréquence conditionnelle.....	119
---	-----

Tableau 3-3. Récapitulatif des principales méthodes de cartographie des matrices de flux	139
--	-----

Conclusion de la première partie –

Tableau C-1. Récapitulatif des problèmes et des solutions envisageables	203
---	-----

Chapitre 7 –

Tableau 7-1. Construction des bornes de classes de distances (Dij) de l'Isère, selon les moyennes emboîtées.....	294
--	-----

Tableau 7-2. Résumé statistique de la matrice (Dij) des liaisons aériennes (Lij).....	315
---	-----

Tableau 7-3. La matrice de flux inter macro régionaux de passagers aériens	320
--	-----

Table des encarts**Chapitre 1 –**

Encart 1-1. La mesure en continu des déplacements individuels	37
---	----

Encart 1-2. Extrait du récit de vie de Marie	40
--	----

Encart 1-3. La période parisienne de Marie	46
--	----

Chapitre 2 –

Encart 2-1. J. Bertin (2000) face à la représentation des matrices	66
--	----

Encart 2-2. Sur la prise en compte du temps en cartographie	80
---	----

Chapitre 3 –

Encart 3-1. Sur le calcul de la migration totale et de la migration nette	125
---	-----

Chapitre 4 –

Encart 4-1. Les lois de la migration (Ravenstein, 1885)	153
---	-----

Encart 4-2. De la physique à la géographie des flux	163
---	-----

Chapitre 5 –

Encart 5-1. Le fichier de détail MOBPRO de l'INSEE.....	223
---	-----

Chapitre 6 –

Encart 6-1. Sur la variation de la mesure de la distance empirique (Eij)	266
--	-----

Encart 6-2. Sur le cheminement de marchandises à l'échelle internationale	268
---	-----

Chapitre 7 –

Encart 7-1. Jules Marey et la segmentation graphique du mouvement	291
---	-----

Chapitre 8 –

Encart 8-1. Exemples de mise en œuvre de la règle des 80/20	337
---	-----

Encart 8-2. Exemples de détermination empiriques du critère de densité de la matrice	339
--	-----

GLOSSAIRE

Les différents termes et expressions sont classés par ordre alphabétique, ceux qui sont portés en *italique* font également l'objet d'une entrée séparée. La signification des notations est présentée dans une table des notations qui suit. La mention d'un astérisque (*) indique que la notion a été introduite dans la thèse, le terme (v. v.) signale un terme lié aux variables visuelles.

Antisymétrie (d'une matrice) : caractère d'une matrice carrée où les valeurs de flux situées de part et d'autre de la diagonale principale sont identiques, mais de signe opposé. Elle est notée $(Fij^-) = (Fij) - (Fij)$, avec $(Fij) = - (Fji)$. Sa cartographie est réalisée par un figuré orienté, une flèche par exemple. Antisymétriser une matrice consiste à prendre la négative de la *transposée* : une matrice (Fij) est *antisymétrique* si ${}^T(Fij) = -(Fij)$.

Asymétrie (des valeurs de flux) : rapport du *solde bilatéral* (Fij^-) sur le *volume bilatéral* (Fij^+) des flux.

Astuce (graphique) : Procédé informel issu de la pratique usuelle permettant de résoudre rapidement un problème posé par la représentation d'un *figuré* sur une carte.

Asymétrie (d'une matrice) : Caractère d'une matrice carrée où les valeurs de flux situées de part et d'autre de la diagonale principale sont, dans l'ensemble, différentes. Elle est notée $(Fij) = (Fij, Fji)$, avec $(Fij) \neq (Fji)$. Sa cartographie est réalisée par un figuré orienté, une flèche par exemple. Nous considérons par défaut que la matrice (Fij) des données initiales (ou observées) est *asymétrique*. Dans la réalité, les flux sont généralement *asymétriques* ; les chances étant, en effet, bien minces pour que la valeur d'un flux soit égale à celle de son reflux.

Bande : correspond à un *Lien* valué, muni d'une *largeur* dont la grandeur illustre la valeur de la relation de *flux* ou de *mouvement* existante entre deux lieux. Elle est caractérisée par ses dimensions géométriques que sont : sa *longueur (carto)graphique*, sa *largeur*, son *périmètre* et sa *surface*. Elle ne se justifie, par opposition au *Lien*, que dans l'unique cas où sa *largeur* est rendue proportionnelle à la valeur de *flux* ou de *mouvement* qu'elle représente (en application de la variable visuelle *Taille*) sinon elle est substitué au *Lien*.

Bilan net (voir *solde*).

Bilatéral (flux) : dans une matrice de *flux asymétrique* ou *antisymétrique*, la valeur de flux est généralement *bilatérale*, ce qui signifie qu'il en existe deux pour un même couple (i,j) de lieux qui présentent des orientations différentes. Il s'agit des valeurs de flux (Fij) et de reflux (Fji) qui sont portées de manière *symétrique* sur les cases situées de part et d'autre de la diagonale principale de la matrice. Elles sont parfois interprétées de manière abusive comme un « aller-retour » entre lieux (i,j), dans la mesure où il n'est pas possible de déduire cette information de la seule information disponible dans la matrice. A noter qu'il n'est pas d'usage d'employer de qualificatif pour désigner des flux unilatéraux, qui ne sont pas *bilatéraux* : on parle de flux tout court.

Composante négative (Fij^-) : sous-matrice (ou composante) *antisymétrique* issue de la décomposition d'une matrice de flux *asymétrique* (Fij) ; cette opération peut être interprétée comme une *symétrisation* : $(Fij^-) = (Fij) - (Fji)$.

Composante positive (F_{ij}^+) : sous-matrice (ou composante) *symétrique* issue de la décomposition d'une matrice de flux *asymétrique* (F_{ij}) ; cette opération peut être interprétée comme une *symétrisation* : $(F_{ij}^+) = (F_{ij}) + (F_{ji})$.

Contiguïté (ou adjacence, connexité) : se dit de deux entités qui sont séparées par une seule limite de zone (ou une seule frontière) de longueur non nulle qu'elles ont en commun. Ces entités sont dites voisines (d'ordre 1) ou limitrophes. Cette contiguïté ordinaire est binaire. Elle peut être ordinale et correspondre au nombre (k) de limites de zones qu'il convient de franchir pour aller d'une entité à une autre, avec (k) variant de 1 à N.

Couleur (v. v) : variable visuelle introduite par J. Bertin (1967) désignant la teinte (ou le ton) d'un objet dont la caractéristique est de type qualitatif nominal (par exemple, un espace agricole est affecté d'une couleur verte). L'utilisation des nuances d'une même teinte, d'un dégradé ou d'un camaïeu, sur différents objets d'une même série de valeurs quantitatives continues correspond à l'application à la variable visuelle *Valeur*, en association avec la *Couleur*.

Diagramme lien-nœud : voir *Graphe*.

Distance cartographique (*) : forme perçue sur la carte de la mesure de la *longueur* d'un *Lien* (ou d'une *Bande*), notée (L_{ij}) et exprimée en millimètres (sur le papier). Forme de distance graphique correspondant à la retranscription de la *distance géographique*, moyennant le rapport d'échelle permettant le passage entre la sphère et le plan.

Distance empirique : distance réelle mesurée sur le terrain, telle que parcourue par l'individu – la route empruntée par le travailleur qui se rend à son lieu de travail ; la voie maritime de la marchandise qui est acheminée vers son aire de marché. Sa cartographie pose ainsi une réelle difficulté liée à la variation de la mesure en fonction des lieux. Aussi est-elle souvent inaccessible ou inconnue et approximée, de ce fait, par la distance traditionnelle à vol d'oiseau (*distance euclidienne*), dont elle diffère pourtant.

Distance euclidienne : distance géométrique mesurée à vol d'oiseau entre deux lieux, indépendamment des contraintes physiques, temporelles et monétaires qui caractérisent le déplacement dans l'espace.

Distance fonctionnelle : coût généralisé de déplacement.

Distance géographique : distance géométrique, exprimée en (kilo)mètres parcourus sur un réseau de transport ou à vol d'oiseau (*distance euclidienne*).

Effet d'alignement topologique des lieux (ou effet de position) (*) : problème général se rapportant à la position relative des lieux, les uns par rapport aux autres. Il dépend de l'échelle d'observation du phénomène. A l'échelle mondiale, cet effet de position se traduit par l'alignement (horizontal ou vertical) des lieux qui empêche la perception claire des flux. Il est particulièrement visible sur les cartes mondiales réalisées sur un planisphère selon une projection Mercator ou Mollweide. A l'échelle locale, cet effet de position se traduit par un *effet d'échelle* lié à la proximité trop importante des couples de lieux entre lesquels, en vertu de la première loi de la géographie de W. Tobler (1969), l'intensité des flux est maximale.

Effet d'amalgame (*) : problème analytique lié au mélange d'informations de flux de natures différentes, qui relèvent ou non de la dynamique du système. Se manifestant également au niveau graphique, par une figure à l'aspect compact. Les solutions apportées à ce problème sont liées au traitement des données, elles consistent par exemple à transformer les valeurs (en fréquences par exemple) et/ou à les décomposer (voir chapitre 3). L'application de solutions sémiologiques à ce problème révèle en fait *l'effet de profondeur*.

Effet d'itinéraire (*) : spécifique à l'échelle mondiale, il résulte de la répartition mondiale terres-mers qui implique de prendre en compte le mode de transport dans la cartographie de certains flux. Par exemple, les flux internationaux de marchandises sont soumis à cet effet dans la mesure où ils sont transportés par

voie maritime, ce qui implique que le tracé des flux contourne les masses continentales. Cet effet est inexistant pour les cartes de mouvements mondiaux qui représentent les trajectoires empruntées par les individus mobiles et, pour certaines cartes de flux réalisées à l'échelle locale.

Effet de convergence : contrainte graphique forte qui consiste à représenter la convergence des flux vers les zones les plus attractives. Sa gestion pratique entraîne une défonction des liens, c'est-à-dire une réduction de leurs dimensions (de la *longueur* des bandes) pouvant être préjudiciable à la lecture de l'information représentée, puisqu'elle entraîne une modification de l'information par rapport à la réalité (une perception des distances parcourues inférieure à la réalité). La *convergence des flux* agit donc sur la *couverture spatiale*, mais pas comme on le souhaiterait.

Effet de couverture spatiale (*) : partie de la complexité graphique d'une carte de flux liée à la dimension géométrique des figurés qui illustrent les flux. Elle se traduit par une densité de teinte perçue trop importante, en raison de *largeurs* trop importantes et de *longueurs* de figurés variées. Le degré de couverture spatiale caractérise la quantité de noir visible sur la figure, par rapport à celle de blanc.

Effet de profondeur (*) : aspect graphique lié à la perception sur la carte, de différentes *longueurs* de figurés. Composante graphique de l'*effet d'amalgame*, il est manifeste sous deux conditions : lorsque la *Transparence* est active et que les figurés sont placés (superposés) en ordre décroissant des *largeurs* des liens (des valeurs de flux) ou bien en ordre croissant de leurs *longueurs*, ce qui revient au même. Cette disposition particulière suggère l'existence d'une composante d'échelle dans la distribution des valeurs de flux, de mobilités qui s'expriment à longue distance ou à une courte distance. Les solutions apportées relèvent donc soit d'une approche sémiologique (chapitre 5), soit théorique (chapitre 6), soit mathématique (chapitre 7).

Effet spaghetti : complexité graphique de l'image des flux, liée à une densité trop importante de figurés présents sur la carte, un rapport visuel de teinte noire sur blanche trop élevé. Résulte de la conjonction d'un *effet de surcharge graphique* et d'un *effet de couverture spatiale*.

Figuré : représentation graphique simplifiée et symbolique d'un objet (par exemple, une ville par un point) ou d'un processus (par exemple, un *flux* par une flèche). Un figuré est caractérisé, d'une part, par ses dimensions géométriques (une bande est définie par sa longueur, sa largeur, son périmètre et sa surface) et d'autre par son esthétique (noir et blanc, en couleur), la teinte de son fond et celle de son contour.

Flux dominant : au sens de Nyusten-Dacey (1961), correspond à la liaison (L_{ij}) entre deux lieux (i) et (j) issue d'un *flux dominant* (soit d'un *flux majeur* émis par un lieu (i) à destination d'un lieu (j) de taille (D_j) plus importante).

Flux interne (ou intra zonal et versus inter zonal) : flux qui ne franchit pas la limite de zone du niveau d'observation (par exemple, à l'échelle communale, le flux qui se produit entre les quartiers d'une même commune est intra zonal)

Flux interzone (versus intra zonal) : se dit d'un flux ayant franchi au moins une limite de zone (flux inter zone ou externe) entre le lieu d'origine et de destination (par exemple, à l'échelle communale, un flux qui va d'une commune à une autre).

Flux majeur : valeur de flux la plus importante, émise par un lieu (i), correspond au flux de rang 1.

Flux (ou flux OD) : mesure d'une quantité (de personnes, de biens, de valeurs, ...) transférée depuis un lieu d'origine vers un lieu de destination, pendant un temps donné [t₁ – t₀]. Il s'agit donc d'une donnée multidimensionnelle qui implique quatre composantes : une thématique (z), deux coordonnées définissant sa position dans un plan (x, y) et une troisième dans le temps (t). Cette donnée est enregistrée dans une *matrice* de flux où l'existence d'une valeur (F_{ij}) correspond au franchissement d'au moins une *limite de zone* entre (i) et (j), le flux est dès lors considéré par défaut comme *interzonal*. La définition du flux ignore le mode de déplacement de la quantité dont il est question, son caractère

matériel ou immatériel. Elle ne prend pas en compte l'espace, lequel n'interviendra donc pas dans le processus de cartographie de flux.

Fond (de carte) : zonage du territoire utilisé pour la collecte, l'analyse et la représentation d'un phénomène. Par exemple, le fond de carte d'une analyse des mobilités régionales est le découpage du territoire en régions.

Fond (de la carte) : aspect thématique de la carte, lié aux données numériques. Par exemple, le fond de la carte qui décrit les résultats d'une analyse des mobilités régionales correspond au thème des déplacements réalisés à l'échelle des régions.

Forme (v. v.) : variable visuelle entraînant la modification de l'apparence graphique d'un objet, en fonction de son implantation. Une modification de *Forme* géométrique appliquée à un point le transforme, par exemple, en carré ou en cercle si le choix de symbolisation est géométrique (elle peut être figurative, abstraite...). La *Forme* permet de différencier les figurés en fonction d'une caractéristique qualitative et non quantitative (qui nécessite la *Taille*).

Graphe (de flux) : dessin de la projection graphique dans un plan quelconque (non nécessairement assorti d'une métrique), sous la forme de points et lignes (de nœuds et d'arcs), de tout ou partie de *l'information disponible* dans une matrice de flux. Le graphe correspond formellement à une structure pauvre car elle est binaire (existence d'un lien ou non), souvent symétrique, généralement antiréflexive et très souvent non spatialisée, à l'inverse de la *carte de flux*.

Graphe (ou diagramme lien-nœud) : terme ambigu désignant à la fois 1) la structure d'un ensemble de sommets et d'arcs archivée dans une *matrice d'adjacence* ; 2) un mode de représentation des structures matricielles, à l'aide de points et de liens et 3) une fonction mathématique. Dans le premier cas, le *graphe* correspond à un réseau de relations sociales, spatiales, ... Dans le second, il fait référence au diagramme lien-nœud auquel il s'est progressivement substitué dans le langage courant. Dans la Théorie des ensembles, on parle de graphe ensembliste pour désigner une structure formée de deux ensembles : celui des sommets et celui des liens, qu'ils soient munis ou non d'attributs particuliers. Ce troisième type de *graphe* désigne donc l'ensemble des couples d'entités qui vérifient une relation donnée, ainsi que la représentation de cette relation sous la forme d'un dessin : le *graphe*.

Graphie : désigne un mode d'écriture dit graphique, sur un support (papier, écran) dont on dirait aujourd'hui qu'il est plutôt vectoriel. Il s'agit d'une manière de communiquer une information, à l'aide du dessin de signes et de caractères conventionnels, le point et la ligne pour le *graphe*, là où les lettres correspondent à l'orthographe, les pas de danse à la chorégraphie.

Information (disponible) : mesure du flux située dans une case de la matrice correspondant à une valeur non nulle. Cette information correspond aussi au taux de remplissage de la matrice, à sa densité exprimée par le nombre de cases renseignées sur le nombre total de cases.

Lien : se dit du *figuré* correspondant à la représentation graphique de l'existence d'une valeur dans la partie centrale de la matrice, autrement dit de l'existence d'une relation entre deux lieux (eux-mêmes étant généralement figurés par des points). Il s'agit du figuré qui illustre l'existence d'une relation, d'une information de *flux* ou de *mouvement*, sans présenter dans valeur. Noté (Lij), il ne présente pas d'épaisseur et n'est défini que par sa *longueur*. Ce lien est symbolisé par un trait, généralement droit, éventuellement orienté quand le flux l'est. Lorsque ce que ce lien est valué, c'est qu'il représente, dans notre cas, un *flux* (Fij) et nous lui substituons le terme de *bande*.

Limite de zone (franchissement de) : correspond à la frontière qui sépare les entités d'origine et de destination. Elle peut être appréhendée comme un point (un poste douanier, par exemple), une ligne ou une surface (périmètre des unités spatiales). Le franchissement d'au moins une limite de zone autorise la mesure d'un flux *inter zonal*, les flux ne franchissant pas cette limite sont qualifiés d'*inter zonal*.

Logique de flux (*) : raisonnement conduisant à la réalisation d'une carte de flux à partir des valeurs quantitatives situées au cœur de la matrice. Sa représentation mobilise une implantation linéaire sous la forme d'un *lien* valué (linéaire ou courbe) dont la *largeur* est proportionnelle à la valeur du flux, en application de la *Taille*. La particularité de cette logique est de ne pas tenir compte de l'espace, à l'inverse de la *logique de mouvement*.

Logique lien-lieu (*) : raisonnement privilégiant une vision des flux scalaire, la représentation des flux du point de vue des lieux, ou de l'effet des flux sur la distribution hiérarchique des lieux. Elle mobilise une implantation généralement ponctuelle et la *Forme* pour différencier les lieux et, éventuellement la *Taille*. Cette logique est qualitative (en raison de la *Forme*) et non purement quantitative comme la *logique de flux* classique.

Logique de mouvement (*) : raisonnement privilégiant une vision relationnelle des flux, tenant compte de l'espace. Elle mobilise un trait sinueux, pour caractériser le cheminement de la quantité déplacée (M_{ij}), de *largeur* proportionnelle (en application de la *Taille*), divisée par la surface de teinte, à la valeur représentée.

Matrice Bertin (ou matrice ordonnée) : type de matrice issue du recodage des valeurs (F_{ij}) en une gamme de couleurs suivi d'une procédure de ré ordonnancement ou de permutation des lignes et des colonnes de manière à obtenir un classement ordonné des valeurs. L'ordonnancement ayant pour objectif de permettre la visualisation de motifs généraux, de réaliser une typologie spatiale.

Matrice (Triangulaire) : sous-partie de la matrice située au dessus (ou en dessous) de la diagonale principale d'une matrice *carrée*. Une matrice *symétrique* peut ainsi être réduite à l'une de ses deux sous-parties triangulaires.

Matrice d'adjacence (ou de connexité) : forme dégradée de *matrice de flux* où la valeur de la cellule est binaire (1,0) et renseigne sur l'existence d'une liaison – ou d'un lien – entre des lieux d'origine et de destination. Correspond à la représentation d'un *graphe* sous une forme matricielle.

Matrice de flux : tableau croisé composé de $(1, \dots, i, \dots, k)$ lignes et de $(1, \dots, j, \dots, p)$ colonnes, correspondant respectivement aux lieux d'origine et de destination des flux et de i, j cellules contenant une valeur quantitative (discrète ou continue). Ce tableau, noté (F_{ij}), correspond aux données observées, mesurées sur le terrain. Elle sera nécessairement fermée, si le nombre de lignes est identique au nombre de colonnes et considérée, par défaut, comme étant *asymétrique*. Composée de N lieux ($N=i=j$), sa diagonale est vide, ce qui porte son nombre de cases renseignées à $N(N-1)$.

Matrice de flux fermée (ou ouverte) : *matrice de flux* (F_{ij}) où les lieux placés en lignes sont identiques à ceux qui sont placés en colonnes ($n*n$), sinon elle est ouverte ($n*m$).

Matrice de flux carrée : *matrice de flux* (F_{ij}) dont le nombre N de lieux d'origine (i) correspond au nombre N de lieu de destination (j), elle est dite de rang N . Cette matrice est *fermée* si les lieux (i) et (j) sont identiques, sinon elle est ouverte.

Matrice dense (versus creuse) : grandes matrices contenant un nombre important d'entités mises en relations, dont la densité de remplissage est importante ou qu'elle est complète (elle comporte peu voire pas de cases vides). Ces matrices denses sont parfois qualifiées de matrices complexes. Elles s'opposent aux matrices *creuses* qui comportent de nombreuses cellules vides (ou dont la densité de remplissage est faible).

Métrique : mesure de l'espacement entre deux entités. Par exemple : la *métrique euclidienne* correspond à la mesure de la distance à vol d'oiseau (en ligne droite) entre deux points.

Mouvement : mesure notée (M_{ij}) d'une quantité transférée depuis un lieu d'origine vers un lieu de destination, localisés sur un espace géographique, pendant un intervalle de temps donné.

Opposition colorée (ou divergence de couleurs) : Principe de mise en œuvre de la variable visuelle *Valeur* appliquée aux données d'une même série statistique, présentant des polarités différentes (+/-), imposant de fait le choix de teintes opposées ou de couleurs divergentes. Les valeurs positives seront, par exemple, teintées d'une couleur chaude et les valeurs négatives, d'une couleur froide.

Représentation (graphique) : issue du latin *repraesentare* (rendre présent) et du grec *graphein* (graphie, écriture), la représentation graphique est une manière de rendre compte visuellement, souvent de manière simplifiée, d'une information. Elle est, dans notre cas, issue d'un tableau de données et figurée à l'aide de symboles conventionnels (des lettres pour un texte) que sont les primitives graphiques : un point, une ligne ou une surface. La représentation graphique des données autorise des allers-retours fructueux entre le tableau, les traitements à réaliser et les résultats obtenus. C'est pourquoi elle possède une place et une importance particulières dans le domaine de l'analyse des données relationnelles, souvent caractérisées par une complexité importante.

Saillance (visuelle) : se dit de la teinte d'un objet dont la luminescence, définie par ses caractéristiques colorimétriques, permet le détachement de son environnement, de « sauter aux yeux ».

Solde de flux : Différence arithmétique marginale (du point de vue des lieux) des valeurs de *flux* entrantes (D_j) et sortantes (O_i) : $(O_i - D_j)$.

Solde bilatéral de flux (ou transfert net) : Différence arithmétique des valeurs de *flux* situées de part et d'autre de la diagonale principale d'une matrice asymétrique, notée $(F_{ij}^-) = (F_{ij}) - (F_{ji})$.

Solde-Solde (*) : type de carte de flux décrivant le *transfert net* du point de vue des liens et des lieux

Surcharge graphique (*) : partie de la complexité graphique d'une carte de flux liée à une densité trop importante de figurés (nombre de liens) qui sont souvent, par ailleurs, fortement superposés et enchevêtrés (traduisant ainsi une *couverture spatiale* trop importante).

Symétrie / Symétrique : Caractère d'une matrice carrée, pour laquelle les valeurs de flux situées de part et d'autre de la diagonale principale sont identiques ($F_{ij} = F_{ji}$), elle est notée (F_{ij}^+) .

Symétrisation (d'une matrice) : réduction d'une matrice *asymétrique* carrée (F_{ij}) en une sous matrice triangulaire (F_{ij}^+) , en appliquant, par exemple, une opération mathématique ou logique. La symétrisation diffère de la transposition $t(F_{ij})$ qui consiste à permuter deux à deux les éléments situés de part et d'autre de la diagonale principale. Une matrice *symétrique* est toutefois égale à sa transposée $(F_{ij}^+) = (F_{ij}) = {}^T(F_{ij}) = (F_{ji})$. Sur le plan graphique, l'opération conduit à diviser par deux le nombre de figurés correspondants aux valeurs renseignées. La matrice qui en résulte est donc *symétrique*.

Taille (v. v.) : variable rétinienne introduite par C.-J. Minard au XIX siècle et formalisée par J. Bertin en 1967, permettant la représentation d'un rapport de proportionnalité entre la valeur du tableau et la surface visuelle du figuré correspondant. Cette transformation permet de différencier les figurés de manière quantitative (et non qualitative, ce qui passe par la Forme).

Transparence (v. v) : variable visuelle introduite par A. Mac Eachren en 1995, sous l'expression « *Value by alpha* » (VBA) pour résoudre le problème de la perception visuelle de données qui sont incertaines ou floues, et celles sont superposées. Appliquée au cas des flux, son utilisation permet de réduire l'*effet de couverture spatiale* en agissant sur la morphologie de la teinte, en l'occurrence sur le canal alpha permettant de gérer l'opacité (ou la transparence) de la teinte.

Valeur (v. v.) : variable rétinienne permettant d'associer à l'une des (k) classes de valeurs quantitatives continues de rapport ou de taux, issues d'une discrétisation cartographique, la nuance correspondante d'une même teinte, également segmentée en (k) paliers. Lorsque la série de valeurs présente des polarités différentes (+/-), il convient d'appliquer le principe de *l'opposition colorée*.

Valeur de flux : mesure du *flux*, observée ou théorique, entre un couple de lieux (i,j). Notée (Fij), elle est archivée dans les cellules portées au croisement des (i) lignes et des (j) colonnes de la matrice.

Variable Visuelle (v. v.) : variable rétinienne, définie par J. Bertin (1967). On distingue généralement les 7 variables suivantes : la *Forme*, le *grain* (texture, structure), la *Taille*, l'*orientation*, la *Couleur*, la *Valeur* et la *localisation* (x, y). La *Transparence* et la *dynamique* font partie des variables généralement ajoutées ces dernières années à cette liste.

Visualisation (graphique) : de données scientifiques – ou visualisation scientifique ou encore visualisation graphique – est une manière de (se) (re)présenter visuellement tout ou partie des données disponibles dans un tableau, par le biais d'une interface graphique. Il s'agit souvent de la représentation simplifiée, d'une réalité souvent complexe, à l'aide de symboles graphiques sur un écran, ce qui correspond à sa principale différence avec la *représentation graphique*.

Voisinage (d'un lieu) : ensemble de ses voisins situé dans un périmètre donné, à une distance (discrète ou continue) donnée, ces derniers pouvant également être rangés de 1 à (N-1) en fonction de leur proximité au lieu (i) d'observation. La relation de voisinage peut être stockée dans une matrice de voisinage, laquelle peut être représentée sur un graphe de voisinage.

Volume de flux : somme arithmétique marginale des valeurs de *flux* entrantes (Dj) et sortantes (Oi) d'un lieu : (Oi+Dj) lieux d'une zone d'étude.

Volume – Volume (*) : type de carte de flux décrivant le *volume de flux* du point de vue des liens (*volume bilatéral*) et des lieux.

Volume – Volume – Asymétrie (*) : type de carte de flux correspondant à une variante du type *volume - volume*, où le symbole proportionnel correspondant au *volume de flux* aux lieux est assorti d'une coloration discrète de l'*asymétrie* (Wij) des valeurs.

NOTATIONS

La définition des termes et expressions portés en *italique* est disponible dans le Glossaire précédent.

- N nombre de lieux de la zone d'étude
- N^2 nombre total de cases de la matrice de flux, formée des N lieux
- $N(N-1)$ nombre total de cases hors diagonale principale, de la matrice de flux formée des N lieux
- n mention d'un lieu quelconque
- (i) lieu d'origine d'un flux (ou d'un mouvement)
- (j) lieu de destination d'un flux (ou d'un mouvement)
- (i,j) entité élémentaire (couple de lieux) décrivant une relation ou un flux
- (x, y) coordonnées géographiques d'un lieu quelconque
- (xi, yi) coordonnées géographiques du lieu d'origine
- (xj, yj) coordonnées géographiques du lieu de destination
- (Pi) région d'appartenance de (i), à un échelon territorial supérieur
- (Pj) région d'appartenance de (j), à un échelon territorial supérieur
- (Oi) Somme marginale de la matrice (Fij) en ligne, correspondant à la somme des valeurs de flux émises par (i), notée : $O_i = \sum_{j=1}^n (F_{ij})$ somme marginale en ligne de (i)
- (Dj) Somme marginale de la matrice (Fij) en colonne, correspond à la somme des valeurs de flux reçues par (j), notée : $D_j = \sum_{i=1}^n (F_{ij})$ somme marginale en colonne de (j)
- (Fij) matrice contenant la valeur du flux mesurée entre (i) et (j). Par convention, (Fij) est considérée comme *asymétrique* (ou non *symétrique*). Elle correspond aux données observées, n'ayant pas fait l'objet d'un traitement poussé, s'opposant ainsi à (Fij)*
- ${}^T(Fij)$ matrice transposée de (Fij) correspondant à la permutation terme à terme des valeurs situées de part et d'autre de la diagonale principale
- (Fij*) matrice (Fij) contenant des données secondaires, issues d'un traitement de type modélisation gravitaire. Il s'agit, en pratique, d'une matrice de type (Fij⁻) où les valeurs positives correspondent à des résidus positifs et celles qui sont négatives, à des résidus négatifs.
- (Fji) matrice contenant la valeur du flux mesurée entre (j) et (i), avec $F_{ji} = {}^T(Fij)$. Par convention, (Fij) étant considérée comme *asymétrique*, (Fji) l'est également, *asymétrique*.
- (Fij⁺) matrice contenant la valeur calculée du *volume bilatéral* de flux entre (i) et (j), issue de la somme des flux entrant et sortants. Par convention, (Fij⁺) est considérée comme *symétrique* ; elle correspond à la *composante* dite *positive* d'une matrice (Fij).
- (Fij⁻) matrice contenant la valeur calculée du *solde bilatéral* de flux entre (i) et (j), ou *transfert net bilatéral*, issue de la différence entre les valeurs de flux entrantes et sortantes. Par convention, on considère que (Fij⁻) est *antisymétrique*, elle correspond à la *composante* dite *négative* d'une matrice (Fij), à sa partie dynamique.
- $Max_i(Fij)$ indique la valeur de flux maximale (ou de rang 1) émise de (i) vers (j), elle correspond au *flux majeur*.
- r(Fij) indique le rang de la matrice, c'est-à-dire le nombre de lignes et de colonnes. Une matrice de rang 3 présente 3 lieux d'origine et de 3 lieux de destination.

-
- (W_{ij}) matrice issue d'un compromis de représentation entre (F_{ij}⁺) et (F_{ji}⁻), illustrée par l'asymétrie de flux, formée par le rapport de (F_{ij}⁻) sur (F_{ji}⁺), $(W_{ij}) = \frac{(F_{ij}^-)}{(F_{ji}^+)}$.
- (V_{ij}) matrice discrète binaire de voisinage de (i) réduite aux k lieux voisins les plus proches de (i) en termes de distance kilo(métriques), avec k=2 à k= (N-1). Surtout utilisée pour le calcul de l'auto corrélation spatiale.
- (A_{ij}) matrice discrète binaire d'appartenance territoriale commune de (i) et de (j), interprétée comme leur situation commune dans une même région ou un même échelon territorial de niveau supérieur (A_{ij})=1, sinon (A_{ij})=0. (A_{ij}) est caractérisée par une certaine continuité spatiale ; aussi appelée connexe, elle est *symétrique* et transitive.
- (C_{ij}) matrice discrète binaire de contiguïté entre (i) et (j), interprétée comme le partage d'une frontière commune de longueur non nulle (C_{ij})=1, sinon (C_{ij})=0. (C_{ij}) est nécessairement *symétrique*.
- (CK_{ij}) matrice discrète de contiguïté ordinale entre (i) et (j), interprétée comme le nombre K de frontières qu'il est nécessaire de franchir pour aller de (i) à (j). Si K=1, alors (CK_{ij})=(C_{ij})=1, sinon (CK_{ij})=1...n. (CK_{ij}) est nécessairement *symétrique*.
- (D_{ij}) matrice de distance géographique (kilo)métrique mesurée entre (i) et (j), généralement sur un réseau de transport (une route par exemple). Elle correspond à une métrique spatiale répondant aux trois propriétés de symétrie (pour tout (i)≠(j), D_{ij}=D_{ji}), de séparation (si D_{ij}=0, alors (i)=(j)) et d'inégalité triangulaire (pour tous (i, j, z), D_{iz} ≤ (D_{ij})+(D_{jz})). En l'absence d'information particulière sur l'éloignement des lieux, (D_{ij}) est souvent assimilée à la distance euclidienne correspondant à la mesure de l'espace à vol d'oiseau.
- (T_{ij}) matrice de distance temporelle, exprimée en unité de temps de parcours pour rejoindre (i) depuis (j)
- (E_{ij}) matrice de coût de déplacement généralisé, exprimée notamment en unité spatiales (kilométriques) et temporelles (de temps) pour rejoindre (i) depuis (j).
- (L_{ij}) Distance cartographique perçue, correspondant à la longueur visuelle d'un figuré représentée sur une carte, généralement exprimée en millimètres sur le papier.

TABLE DES MATIERES

INTRODUCTION GENERALE.....	13
GENÈSE ET PROBLÉMATISATION DU SUJET	15
DÉMARCHE MISE EN ŒUVRE	20
CADRAGE THÉORIQUE DU SUJET	22
PLAN GÉNÉRAL	24
PREMIERE PARTIE ETAT DE L’ART DES MÉTHODES DE CARTOGRAPHIE DE FLUX. ÉNONCÉ DES PROBLÈMES, DES CONTRAINTES GRAPHIQUES ET DES DIFFICULTÉS D’ORDRE MÉTHODOLOGIQUE.....	25
INTRODUCTION DE LA PREMIÈRE PARTIE	25
1 ^{ER} CHAPITRE. LA CONSTRUCTION DE LA MATRICE DE FLUX : INCERTITUDES DE MESURE ET PROBLÈMES D’AGRÉGATION	27
INTRODUCTION DU CHAPITRE	27
1.1 Présentation de la matrice de flux	28
1.1.1 Structure et propriété générale de la matrice de flux.....	29
1.1.2 La valeur (Fij) des cellules hors diagonale de la matrice	31
1.2 La mesure du flux : une double agrégation spatio-temporelle	35
1.3 La construction de la matrice de flux : l’exemple des migrations humaines	39
1.3.1 Éléments sur l’approche compréhensive de la construction de la matrice de flux que nous proposons.....	40
1.3.2 La procédure de filtrage de la mesure	44
1.3.3 La procédure d’agrégation spatiale des « paquets ».....	48
CONCLUSION DU CHAPITRE	54
2 ^{ÈME} CHAPITRE. LES MODÈLES DE LA REPRÉSENTATION GRAPHIQUE DES MATRICES DE FLUX : LA MATRICE ORDONNÉE, LE GRAPHE ET LA CARTE	55
INTRODUCTION DU CHAPITRE	55
2.1 La matrice ordonnée.....	58
2.1.1 L’objectif.....	58
2.1.2 Le principe.....	60
2.1.3 La contrainte.....	60
2.1.4 Les variantes.....	61
2.2 Le diagramme liens-nœuds ou graphe	66
2.2.1 L’objectif.....	67
2.2.2 Le principe.....	70

2.2.3 Les contraintes.....	71
2.2.4 Les variantes.....	76
2.3 La carte de flux	77
2.3.1 L'objectif.....	79
2.3.2 Le principe.....	85
2.3.3 Les contraintes.....	86
2.3.4 Les variantes.....	93
CONCLUSION DU CHAPITRE	106
3 ^{IÈME} CHAPITRE. LA DÉCOMPOSITION DES MATRICES DE FLUX (Fij) : LE VOLUME (Fij ⁺) ET LE SOLDE (Fij ⁻) BILATÉRAL.....	109
INTRODUCTION DU CHAPITRE	109
3.1 <i>L'information disponible</i> dans la matrice : ce que l'on peut cartographier (Lij)	111
3.1.1 L'information disponible dans la matrice.....	111
3.1.2 Des liens et des lieux : l'analogie de la matrice de flux avec le tableau de contingence... 115	
3.2 Les approches statistiques et matricielles du traitement des données de flux (Fij).....	117
3.2.1 L'approche statistique : la cartographie des probabilités de fréquence des valeurs de flux 117	
3.2.2 L'approche matricielle	120
3.2.3 L'approche de W. Tobler : l'analyse des composantes symétriques (Fij ⁺) et antisymétriques (Fij ⁻) de la matrice.....	124
3.3 Les trois grandes méthodes de cartographie des matrices de flux	127
3.3.1 La cartographie des matrices de flux asymétriques : l'exemple des valeurs observées (Fij) 128	
3.3.2 La cartographie des matrices de flux symétriques, l'exemple des volumes de flux bilatéraux (Fij ⁺) 129	
3.3.3 La cartographie des matrices de flux antisymétriques : l'exemple des transferts nets bilatéraux (Fij ⁻)	131
3.3.4 La cartographie des matrices de flux issues d'une combinaison entre (Fij ⁺) et (Fij ⁻), l'exemple de l'asymétrie (Wij).....	135
CONCLUSION DU CHAPITRE	139
4 ^{IÈME} CHAPITRE. L'ÉVOLUTION DES MÉTHODES DE CARTOGRAPHIE DE FLUX : ÉTAT DE L'ART DES PROCÉDÉS ET DES PRINCIPES DE SÉMIOLOGIE GRAPHIQUE	141
INTRODUCTION DU CHAPITRE	141
4.1 Rétrospective (1786 -) des méthodes de cartographie de flux	143
4.1.1 L'origine (H. D. Harness, 1836).....	145
4.1.2 L'âge d'or (C.-J. Minard, 1844-1877).....	146
4.1.3 Le tournant (E.G. Ravenstein, 1885).....	153

4.1.4 La vision globale (E. Raisz ; J. Bartholomew, 1920- ; Renner, 1942), l'ère de « la géographie aérienne »	158
4.1.5 L'apogée (C. W. Thorntwaite, T. Hägerstrand, W. Tobler, 1934-1987).....	162
4.1.6 A l'heure de l'Informatique graphique : vers le renouvellement de la sémiologie	169
4.2 La conception d'une carte de flux : aspects graphiques et sémiologiques.....	177
4.2.1 Aspects sémiologiques de la cartographie de flux	179
4.2.2 Du tracé à la disposition du figuré.....	187
4.2.3 Les contraintes de la flèche	189
4.2.4 Les variantes de la flèche	194
CONCLUSION DU CHAPITRE	198
CONCLUSION DE LA PREMIÈRE PARTIE	201
DEUXIEME PARTIE PROPOSITIONS DE SOLUTIONS VISANT À RÉDUIRE LA COMPLEXITÉ DE LA CARTE DE FLUX.....	205
INTRODUCTION DE LA DEUXIÈME PARTIE	205
5 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES AUX ASPECTS SÉMIOLOGIQUES ET ESTHÉTIQUES DE LA CARTE DE FLUX : LE RENOUVELLEMENT DE LA VARIABLE VISUELLE <i>COULEUR</i>	211
INTRODUCTION DU CHAPITRE	211
5.1 La nouvelle esthétique de la carte de flux	213
5.1.1 Le choix de la couleur	214
5.1.2 De l'intérêt d'un fond sombre	216
5.2 L'introduction de la variable visuelle <i>Transparence</i> , l'exemple du (Fij ⁺).....	222
5.2.1 Présentation des données.....	223
5.2.2 Le canal Alpha de la teinte	224
5.2.3 Application de la <i>Transparence</i> à la cartographie du volume bilatéral (Fij ⁺).....	225
5.2.4 Expérimentation sur l'intérêt esthétique et l'efficacité du fond noir et de la <i>Transparence</i> 227	
5.3 Le renouvellement de la variable visuelle <i>Couleur</i> , l'exemple du (Fij ⁻).....	231
5.3.1 La sémantique de l'opposition colorée.....	231
5.3.2 Application empirique de l'opposition colorée à la cartographie du solde bilatéral (Fij ⁻)	233
5.4 Validation empirique : l'exemple de flux commerciaux internationaux.....	235
5.4.1 Présentation des données.....	235
5.4.2 Les solutions proposées.....	237
CONCLUSION DU CHAPITRE	246

6 ^{IÈME} CHAPITRE. D'UNE CARTOGRAPHIE DE FLUX (FIJ) À UNE CARTOGRAPHIE DE MOUVEMENTS (MIJ) ET INVERSEMENT	247
INTRODUCTION DU CHAPITRE	247
6.1 La réalisation d'une carte de flux selon une <i>logique de flux</i> (Fij).....	250
6.1.1 L'hypothèse n°1 d'une perception des variations de 'largeur' du figuré : un paradoxe....	254
6.1.2 L'hypothèse n°2 d'une perception des variations de la 'surface visuelle' du figuré.....	256
6.2 Les distances (Dij) mobilisables en cartographie de flux (Fij)	258
6.2.1 La distance cartographique (Lij)	258
6.2.2 La distance géométrique : géographique ou euclidienne (Dij).....	263
6.2.3 La distance empirique (Eij)	266
6.3 D'une <i>logique de flux</i> (Fij) à une <i>logique de mouvement</i> (Mij)	272
6.3.1 Définition de la quantité de mouvement (Mij)	272
6.3.2 L'hypothèse n°3 d'une perception du 'mouvement'	275
6.3.3 L'hypothèse n°4 d'une perception 'approximative du mouvement' sur une carte de flux	278
CONCLUSION DU CHAPITRE	282
7 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES À LA PRISE EN COMPTE DE L'ESPACE : LA DISTANCE, LE VOISINAGE SPATIAL ET L'APPARTENANCE TERRITORIALE	285
INTRODUCTION DU CHAPITRE	285
7.1 La métrique réelle ou <i>distance géographique parcourue</i> par le flux (Dij)	287
7.1.1 L'approche sémiologique : le recours à la Transparence	288
7.1.2 L'approche cartographique: la collection de cartes, issues d'une segmentation de la matrice	290
7.2 La métrique de proximité : le <i>voisinage</i> spatial binaire (Cij) ou ordinal (CKij).....	299
7.2.1 La contiguïté simple (Cij=1)	300
7.2.2 La contiguïté ordinale (CKij)	302
7.3 La similarité des lieux : le <i>critère d'appartenance territoriale</i> [Fij ← (Pi=Pj)].....	306
7.4 Validation empirique : l'exemple de flux aériens internationaux de passagers	311
7.4.1 Présentation des données.....	311
7.4.2 Cartographie des vols aériens locaux : le « petit courrier » (Fij) ← (Dij<Q1).....	314
7.4.3 Cartographie des flux aériens entre pays limitrophes : le petit courrier entre voisins (Fij) ← (Cij=1)	315
7.4.4 Cartographie des flux aériens (inter)régionaux : les moyen et long courriers (Fij) ← (Pi=Pj)	317
CONCLUSION DU CHAPITRE	323

8 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES AU CHOIX D'UN CRITÈRE DE SÉLECTION GLOBAL : LA CARTOGRAPHIE DE <i>FLUX SIGNIFICATIFS</i> ($F_{ij} > \alpha$)	325
INTRODUCTION DU CHAPITRE	325
8.1 Le critère ($F_{ij} \geq \alpha$) de la sélection directe des flux à cartographier.....	327
8.1.1 La cartographie du Top-n des flux	327
8.1.2 Le critère ($F_{ij} \geq \alpha$) : de l'examen de la forme de la distribution des valeurs de flux à l'étude de leur concentration.....	328
8.2 Le critère de concentration (Lorenz) des fréquences de valeurs de flux : (F_{ij} vs L_{ij}).....	332
8.3 L'estimation du double critère de sélection ($F_{ij}, L_{ij} > \alpha$) de la part (en %) de l'information de flux en fonction de la part de liens représentée et inversement	335
8.3.1 L'objectif de significativité des valeurs ($\alpha \approx 80\%$).....	336
8.3.2 L'objectif de densité des valeurs : $(2 - 5) \%$	338
8.4 Vers le développement d'outils informatiques : suggestion de programmes pour Rstudio© ...	340
8.4.1 Vers le développement d'une interface interactive	341
8.4.2 Suggestion de programmes RStudio©, de sélection de flux à cartographier selon un critère global	343
CONCLUSION DU CHAPITRE	345
9 ^{IÈME} CHAPITRE. PROPOSITIONS DE SOLUTIONS LIÉES AU CHOIX D'UN CRITÈRE DE SÉLECTION LOCAL : LA CARTOGRAPHIE DE <i>FLUX DOMINANTS</i>	347
INTRODUCTION DU CHAPITRE	347
9.1 La méthode Nyusten-Dacey (1961) d'analyse des <i>flux dominants</i>	350
9.1.1 Présentation des données (de la matrice disponible pour l'analyse)	350
9.1.2 Représentation graphique des relations nodales.....	352
9.1.3 Cartographie de la structure nodale et régionalisation	353
9.2 Proposition de reconstruction cartographique de la méthode Nyusten-Dacey (1961)	357
9.2.1 Etape 1 : cartographie des flux majeurs $F_{im} = \text{Max}_i(F_{ij})$	358
9.2.2 Etape 2 : d'une logique de flux à une logique de liens-lieux	359
9.2.3 Etape 3 : cartographie des « flux dominants »	362
9.2.4 Etapes 4 et 5 : régions et sous-régions polarisées.....	364
9.2.5 Cartographie des flux dominants : l'exemple des navettes de l'Isère (F_{ij})	367
9.3 Validation empirique : l'exemple de flux commerciaux internationaux (OMC, 2007)	374
9.3.1 Cartographie des flux commerciaux dominants (F_{ij}), le cas officiel.....	375
9.3.2 Le cas légal : la cartographie des flux dominants (F_{ji}).....	386
9.3.3 Le cas évident : la cartographie des flux dominants (F_{ij+})	389
CONCLUSION DU CHAPITRE	392
CONCLUSION DE LA DEUXIÈME PARTIE.....	395

CONCLUSION GENERALE	399
PRINCIPAUX APPORTS DE LA THÈSE	400
PERSPECTIVES.....	405
BIBLIOGRAPHIE.....	1
ARTICLES ET OUVRAGES.....	1
DOCUMENTS CARTOGRAPHIQUES	15
RÉFÉRENCES TECHNIQUES	17
RÉFÉRENCES PERSONNELLES	20
LISTE DES TABLES.....	25
TABLE DES FIGURES.....	25
TABLE DES TABLEAUX.....	32
TABLE DES ENCARTS.....	32
GLOSSAIRE	33
NOTATIONS.....	41
TABLE DES MATIERES.....	43
ANNEXES.....	1
SOMMAIRE DES ANNEXES.....	3
LISTE DES ANNEXES DES CHAPITRES.....	3
LISTE DES PROGRAMMES POUR RSTUDIO© ANNEXÉS.....	3

ANNEXES

« La recherche en visualisation est impossible sans le développement de prototypes de logiciels permettant l'élaboration et la validation des idées »²²⁶

Natalia et Gennady Adrienko.

²²⁶ Texte traduit par nos soins, la citation exacte est la suivante : « *Visualisation research is impossible without developing software prototypes for checking and elaborating ideas* ».

SOMMAIRE DES ANNEXES

Les annexes sont numérotées selon leur ordre d'apparition dans le texte.

Liste des annexes des chapitres

Annexe A.1. Les méthodes de collecte des flux migratoires et commerciaux

Annexe A.2 La construction d'une matrice de voisinage spatial (k plus proches voisins)

Annexe A.3. La transformation d'une matrice ordinaire (F_{ij}) en une matrice ordinale

Liste des programmes pour Rstudio© annexés

Annexe R-1 : Procédure de transformation d'une matrice de flux fermée non carrée en une matrice de flux fermée carrée

Annexe R-2 : Procédure de cartographie d'une matrice de flux sous RStudio (approche matricielle)

Annexe R-3 : Procédure de sélection des flux à cartographier selon un critère spatial (de distance parcourue ou de voisinage)

Annexe R-4 : Procédure de sélection des flux à cartographier, selon un critère global

Annexe R-5 : Procédure de sélection des flux à cartographier, selon un critère local (flux dominants)

Annexe R-6 : Exemple de cartographie directe d'une variable (ou matrice) de l'environnement global de Rstudio©

ANNEXES DES CHAPITRES

Annexe A. 1. Les méthodes de collecte des flux migratoires et commerciaux

Cette annexe qui porte sur la mesure des flux concerne le chapitre 1.

A.1.1 La collecte des flux migratoires

Deux grands groupes de méthodes sont mis en œuvre pour collecter des flux migratoires et les dénombrer de manière plus ou moins précise : le registre de population et le groupe des enquêtes et des sondages.

La méthode directe de mesure des déplacements humains la plus complète est le registre de population. Il s'agit d'un « *Système de collecte des données par les services publics selon lequel les caractères démographiques et socioéconomiques de l'ensemble d'une population ou d'une partie de la population font l'objet d'un enregistrement continu* »²²⁷. Dans un registre de population, c'est la donnée qui porte sur les changements de résidence de l'individu entre deux dates qui permet de reconstituer « *le cheminement migratoire* » (Poulain, 1981 : 18), c'est-à-dire la trajectoire spatiale d'un individu. Le registre permet donc d'identifier avec précision, les individus qui ont changé de lieu de résidence et qui sont considérés comme des migrants. Lorsque le fichier est centralisé, les effectifs de migrants sont estimés en comparant la localisation des lieux de résidence de chacun d'entre eux, en début et en fin de période. Ces registres, qui permettent de réaliser des analyses longitudinales et des analyses transversales, n'existent malheureusement que dans quelques pays : ils n'existent pas en France, par exemple. Dans les pays qui en ont mis en place, l'information collectée est toutefois limitée à un territoire précis, en raison de la lourdeur et du coût de la procédure à mettre en œuvre. Le registre de population est donc un cas bien particulier de collecte des données qu'il convient de distinguer du second groupe de méthodes.

Le second dispositif de mesure directe des déplacements regroupe les enquêtes et les sondages. Il est organisé autour de deux grands types d'outils : l'enquête rétrospective et l'enquête à passages répétés.

L'*enquête rétrospective* permet de connaître *a posteriori*, les positions successives d'un individu n en différents lieux i et les dates associées. La mesure porte, par exemple, sur les changements de résidence, déclarée comme principale, entre deux dates, pour un échantillon représentatif de la population. Elle pourrait aussi bien porter sur des lieux de travail successifs ou des lieux de loisirs. L'enquête rétrospective peut être mise en œuvre à différentes échelles de la vie d'un individu : à celle d'une vie pour saisir l'ensemble de ses migrations ou à celle de ses déplacements quotidiens, pour saisir sa mobilité domicile-travail, par exemple. Dans le premier cas, il s'agit d'enquêtes biographiques tandis que dans le second cas, il est question d'enquêtes spécifiques sur les mobilités quotidiennes. Celles-ci sont constituées en interrogeant une personne par ménage (ou l'ensemble des personnes par ménage) sur ses (leurs) propres déplacements réalisés la veille, le dernier week-end ou durant la semaine de l'enquête ainsi que, notamment sur ses (leurs) motifs²²⁸ du déplacement qui définissent le type de mobilité, sur ses (leurs) modes de transports. L'interrogation d'une personne par ménage est mise en œuvre dans l'Enquête Transport réalisée conjointement par l'INSEE et l'Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (IFSTTAR, ex - Institut National de Recherche sur les Transports et leur Sécurité, INRETS). L'interrogation de

²²⁷ Cette définition est disponible sur le site de l'Institut National d'Études Démographiques (INED), à l'adresse suivante : <http://www.ined.fr/fr/lexique/bdd/mot/Registre+de+population/motid/82/>

²²⁸ Les motifs de déplacement quotidiens sont, par exemple, les trajets « Domicile-travail » ; « Domicile-études » ; « Domicile-achats » ; « Domicile-affaires personnelles » ; « Domicile-accompagnement » ou « Domicile-loisirs » qui sont en lien avec le lieu de résidence (le domicile). Lorsqu'ils ne sont pas en lien, les déplacements sont généralement considérés comme secondaires.

l'ensemble des personnes du ménage sur leurs déplacements personnels réalisés la veille du jour de l'enquête est mise en œuvre dans les enquêtes Enquêtes-Ménages Déplacements (EMD) menées par le Centre d'Etudes sur les Réseaux, les Transports, l'Urbanisme et les Constructions (CERTU) dans certaines villes françaises (voir Encart A9-1), ainsi que dans les enquêtes de la Direction Régionale de l'Équipement de la Région Ile-de-France (DREIF) pour la région parisienne. D'autres méthodes²²⁹ de collecte des déplacements individuels sont mises en œuvre dans le champ des transports, parmi lesquelles l'Enquête Déplacements Villes Moyennes et l'EMD sur de grands territoires (CERTU), l'Enquête Nationale Transports (INSEE-IFSTTAR ex. INRETS) et l'Enquête Globale Transports (EGT) du Syndicat des Transport de l'Ile-de-France (STIF) et de la DRIEA pour la région parisienne.

Encart A9-1. L'Enquête Ménage Déplacement du CERTU

L'enquête rétrospective typique sur les déplacements urbains quotidiens en France est l'EMD qualifiée de « méthode standard CERTU » (CERTU, 2008). Celle-ci porte sur la collecte des déplacements individuels quotidiens de la veille du jour de l'enquête, de l'ensemble des membres d'un ménage de plus de 5 ans. Le questionnaire d'enquête de l'EMD est composé de cinq modules : *i*) un module sur le déplacement ; *ii*) un autre sur le ménage ; *iii*) un troisième sur l'individu. *iv*) Le quatrième module porte sur le trajet, en lien avec le déplacement. Un déplacement est donc composé de plusieurs trajets dès lors que le mode du déplacement change. *v*) Le cinquième module porte sur l'opinion des individus sur des actions urbaines particulières et spécifiques à l'agglomération sur laquelle porte l'enquête. Les données sont collectées sur deux maillages différents qui constituent autant de filtres territoriaux de la mesure (voir 1.3.2) : le premier est celui des zones de tirages au sein desquelles le nombre de ménages enquêté est constant. Le second est celui des zones d'origine et de destination dont les limites recouvrent partiellement celles des agglomérations, des communes ou des découpages infra-communales. Les EMD permettent donc de connaître la situation des individus à l'échelle des zones par tranche horaire, qui est le filtre temporel retenu. Elles sont souvent analysées conjointement avec les Enquêtes Emplois du Temps (EDT)²³⁰ de l'INSEE qui collectent des données plus spécifiques sur les motifs des déplacements des individus de 15 ans et plus, la veille du jour de passage de l'enquêteur. Les activités des individus sont ainsi reconstituées, de manière rétrospective, par pas de temps de dix minutes. Elles sont ensuite recodées selon une nomenclature particulière sur la base de « temps sociaux ». L'intérêt de l'EDT est la prise en compte du temps dans le déplacement, par l'introduction de la notion de « Budget-temps ». Son principal inconvénient, de notre point de vue, est l'absence de référence spatiale.

Aux côtés des enquêtes rétrospectives peuvent être mises en œuvre des enquêtes à passages répétés, afin d'observer les événements démographiques. Celles-ci sont réalisées à partir du passage d'un questionnaire à intervalles réguliers, dans une même zone d'étude, de 6 mois ou de 1 an qui correspond aux normes de la migration définitive. Elles permettent d'enregistrer les entrées et les sorties de chaque unité spatiale au cours des intervalles de temps. Si les mouvements sont bien collectés, le lieu de destination des migrants n'est pas connu, l'information collectée sur les déplacements est ainsi lacunaire. Ce type d'enquête est surtout utilisé dans les pays où il n'existe pas de source statistique fiable.

²²⁹ Une présentation synthétique de l'ensemble des méthodes de collecte de données sur les déplacements individuels et les transports, mises en œuvre par le CERTU, est disponible dans : CERTU, 2008, Note de synthèse « Mobilités et transports », *Les recueils de données sur la mobilité urbaine*, Hors collection, n° 2, 8 p.

²³⁰ La dernière version de cette enquête date de 1998-1999.

Le troisième dispositif de mesure directe des déplacements est celui des *Recensements de la Population* (RP) réalisés par l'INSEE. Les RP constituent une source de données très utilisée. Ils permettent la collecte d'une information à une date précise (t) qui porte, soit sur une date antérieure : à (t_1), soit sur un intervalle (z) de temps $[t_1-t_2]$ à partir de la réponse à une question unique qui prend par exemple, les formes respectives suivantes pour la collecte spécifique des migrations résidentielles : « Où habitez-vous en 1995 ? »²³¹ ou bien « Où habitez-vous il y a 5 ans ? »²³² et « Vous êtes-vous déplacé au cours des six derniers mois ? ».

Les résultats diffèrent selon la formulation de cette question. Lorsque qu'elle porte sur le lieu de résidence à une date antérieure, la réponse de l'enquêté permet de déterminer le nombre de migrants²³³ en comparant la localisation de leur résidence au jour de l'enquête, à celle de la date antérieure ; elle permet aussi de dénombrer au moins un déplacement dans l'intervalle de temps considéré : les migrations de retour et les décès sont donc exclus de cette mesure. La question sur la localisation du lieu de naissance par rapport au lieu de résidence au jour de l'enquête renvoie également une seule migration par migrant, si la localisation des deux lieux diffère. La question sur le lieu de résidence à la date d'un précédent recensement permet de ne saisir que le dernier déplacement, c'est-à-dire la dernière migration. Enfin, la question qui porte sur la durée de présence dans la résidence actuelle est identique à la précédente, elle apporte en sus une dimension temporelle à la mesure de la dernière migration.

La réponse unique, à une question unique, caractéristique des recensements, ne peut généralement renvoyer qu'un seul déplacement, c'est-à-dire qu'une seule migration pour un individu. Or, la probabilité de changer de lieu de résidence entre deux dates de recensement est plus ou moins importante selon les périodes de la vie. Il est, en outre, fort probable que certains individus aient effectués plusieurs migrations (Courgeau, 1975a), qui ne sont donc pas comptabilisées. Les recensements sont, par conséquent, une source d'information très sensible²³⁴ au choix des référentiels²³⁵ temporels et spatiaux utilisés lors de la collecte. C'est pourquoi la section 1.3 (page 39) dédiée à la construction de la matrice des flux s'appuiera sur la présentation de cette méthode.

La mesure directe des effectifs d'individus qui se sont déplacés sur une période entre deux lieux, celle des migrants, tout comme celle de leurs déplacements, les migrations, est par essence incomplète ou sous-estimée. Cette situation a conduit au développement de méthodes de collecte indirectes qui consistent à reconstituer des flux à partir de stocks. C'est le cas, par exemple, des données qui portent sur des flux internationaux d'étrangers ou de réfugiés collectées par l'Organisation Internationale des Migrations (OIM), sur la base d'une déclaration des Etats. Outre l'incertitude de l'information collectée sur la base d'une déclaration, les données renseignent sur le nombre de résidents de nationalité étrangère par pays, c'est-à-dire sur un effectif d'individus et non pas sur un flux. C'est ainsi que l'existence d'une relation entre des pays d'origine et des pays d'accueil de migrants, qui est appréhendée en tant que structure relationnelle évaluée, est analysée en termes de flux. Des méthodes alternatives sont aussi mises en œuvre : elles recourent à des marqueurs à partir desquels sont reconstitués des déplacements. C'est le cas des fichiers d'abonnements et de désabonnements à un fournisseur d'énergie (électricité, eau), de télécommunications ou des fichiers électoraux qui sont utilisés pour reconstituer des mobilités résidentielles.

²³¹ Cette première formulation correspond aux recensements réalisés entre 1962 et 1999, pour lesquels la question portait sur la localisation du lieu de résidence au premier janvier de l'année du précédent recensement.

²³² Cette seconde formulation correspond aux nouveaux recensements ou recensement rénovés réalisés depuis 2004.

²³³ La méthode ne permet pas de déterminer directement le nombre de migrations.

²³⁴ Ils sont par ailleurs très sensibles à l'effet-mémoire des répondants qui a une influence sur le degré d'erreur et la précision des données collectées, par conséquent sur leur fiabilité.

²³⁵ Il est nécessaire de tenir compte des filtres temporels et territoriaux pris en compte dans la mesure.

A.1.2 La collecte des flux commerciaux

La mesure discrète des flux de marchandises suit une toute autre logique que celle des flux de populations, car elle est nécessairement appréhendée d'une part, par le biais de leur transport²³⁶ dans des véhicules, qui caractérisent le mode de transport et d'autre part, sur une distance exprimée en kilomètres, en temps ou en valeur de déplacement. Tout cela contribue, en effet, à caractériser la valeur de la marchandise transportée. Le flux commercial est, par conséquent, d'abord soumis à des impératifs économiques et sociaux de consommation qui incluent la demande de transport. Il est aussi dépendant d'impératifs de production qui intègrent l'offre des transporteurs. La mobilité spatiale des biens, à l'inverse de celle des personnes, s'intègre par ailleurs dans un dispositif complexe formé par l'articulation de nombreux facteurs qui forment un système et qui sont souvent spécifiques à une filière particulière (le transport de fleurs, par exemple) et à un mode de transport particulier (ferroviaire, routier, maritime, fluvial ou aérien).

Afin de rester dans le cadre général de constitution d'une matrice de flux, nous appréhendons dans un premier temps, ces flux commerciaux indépendamment du mode de transport utilisé et de la filière à laquelle ils appartiennent, c'est-à-dire indépendamment du type de produit. Nous restons dans le seul champ dit de la « consommation du transport », celui des expéditeurs, aussi appelés les « chargeurs » (du côté de la demande de transport) et non dans celui des producteurs de transport définis comme les propriétaires des moyens de transports (du côté de l'offre de transport).

La seconde particularité de la mesure des flux de marchandises est qu'elle est toujours soumise à une unité de mesure exprimée en tonnes, ou tonnes.kilomètre. Le volume ainsi mesuré permet d'exprimer la distance parcourue par la marchandise qui est une notion primordiale dans l'analyse de la mobilité des biens. L'utilisation du tonnage kilométrique est privilégiée dans de nombreuses sources de données car il permet, d'après M. Savy, de ne pas faire varier l'unité de la marchandise en fonction des caractéristiques de la trajectoire, selon que le trajet est direct ou qu'il comporte des lieux intermédiaires de transit, dits de rupture de charge. On notera que le volume en m³ ou la taille ne sont pas considérés comme des critères caractéristiques des flux de marchandises, alors qu'ils correspondent aux critères les plus pertinents (Savy, 2007 : 56-57). D'après l'auteur, le volume d'une marchandise exprimée en tonnes n'est pas révélateur de la réalité du marché économique et de la prestation correspondante, il donnerait même une vision déformée de la réalité, à l'inverse des unités monétaires supposées mieux exprimer l'activité économique.

Le prix du transport est, en effet, un indicateur courant du transport de biens, il varie en fonction du type de marchandise (de sa valeur économique), de son mode de transport et du temps de déplacement, *à fortiori* de la trajectoire empruntée. Si l'on considère à titre d'exemple, l'organisation mondiale des flux commerciaux, on observe que les biens les moins onéreux en valeur monétaire sont transportés, par exemple depuis l'Asie du sud-est vers l'Europe, par voie maritime et non aérienne, à l'instar de ceux qui ont le plus de valeur ; le temps de transport en conteneurs maritimes est, de fait allongé, par rapport à celui du fret aérien, tout comme la distance parcourue est plus importante dans le premier cas.

Les flux commerciaux, tout comme les flux migratoires, se distinguent en fonction de l'échelle spatiale à laquelle ils se produisent, échelle qualifiée de *nature du transport*. On différencie ainsi le *transport intérieur* (à un pays) du *transport international* réalisé entre différents pays, comme on distingue les *migrations internes* des *migrations internationales*. A noter que pour les flux migratoires, le terme de migration est généralement attribué dans l'usage courant aux migrations internationales et celui de mobilité spatiale aux migrations internes.

²³⁶ Nous avons vu que les flux de populations pouvaient également être appréhendés par le biais de leur mode de transport (par voie ferroviaire, par exemple).

Les flux commerciaux présentent la particularité d'être décomposables en quatre types de flux, à l'inverse des flux migratoires. Est, en effet, distingué pour les flux commerciaux internationaux un *transport de transit* qui correspond aux flux traversant des unités spatiales, sans s'y arrêter : celles-ci ne sont, de fait, ni l'origine ni la destination finale de la marchandise.

Le flux de transit diffère de *l'échange*, défini comme le transport international dont l'origine ou la destination de la marchandise sont situées en France. Cette notion de transit est étendue à d'autres échelles. Ainsi, le *transit inter régional* concerne, par exemple, le transport de marchandises dont les lieux d'origine, ceux de *chargement* et les lieux de destination, ceux de *déchargement* de la marchandise, ne sont pas situés dans les régions dont il est question. Une matrice des flux commerciaux peut, par conséquent, être observée du point de vue de flux entrants ; de flux sortants ; de flux bilatéraux ainsi que de flux de transit. Il convient toutefois de noter que ces flux de transit font l'objet d'une seconde matrice, les données étant collectées dans le cadre d'enquêtes spécifiques. L'Enquête *Transit*, par exemple, est réalisée auprès des poids lourds qui franchissent des points particuliers tels que les Alpes. D'après le Service de l'Observation et des Statistiques (SOEs)²³⁷ du Ministère français de l'écologie, du développement durable et de l'énergie (MEDDE), pour l'année 2013, l'enquête *Transit* « [...] permet de connaître tous les 5 ans l'évolution de la structure par origine/destination et nature des marchandises notamment des trafics dont un suivi est assuré par ailleurs (bulletin *Alpinfo* édité par la Suisse avec le concours de la France et de l'Autriche sur l'ensemble de l'arc alpin, publications annuelles de l'Observatoire franco-espagnol des trafics à travers les Pyrénées) ».

Une autre particularité de la mesure des flux de marchandises, par rapport aux flux migratoires, est qu'elle est parfois réalisée en date réelle, ce qui signifie qu'elle est actualisée au jour le jour en des lieux particuliers. Les flux estimés dans le champ des transports sont, en général, appréhendés au regard d'une infrastructure formée par un réseau (routier, ferroviaire). Cette infrastructure est composée de nœuds et d'arcs, elle forme le réseau de transport proprement dit, c'est-à-dire les voies de communication (Savy, 2007). La mesure des flux est dès lors disponible au niveau des arcs ou bien à celui des nœuds. Les nœuds du transport par voie routière correspondent aux lieux d'observation des mouvements de marchandises, aux entrepôts et plateformes de stockage. L'information est donc fondée non pas sur la base d'un déplacement entre deux lieux (d'origine et de destination) mais sur le flux (d'entrée et de sortie) observé en un lieu particulier.

A l'instar des flux migratoires, les méthodes de mesure des flux commerciaux sont, de manière générale, fondées sur le principe d'une collecte rétrospective réalisée par mode de transport et par nature de produit ou bien sur une collecte prospective dans le cadre du transport international.

Le Système d'information sur les transports de Marchandises (SitraM), mis en place par la France en 1974, fournit les données les plus complètes sur les flux de marchandises observés du point de vue des consommateurs. SitraM, réalisé par le SOeS et actualisé tous les ans, rassemble différents fichiers de données numériques qui portent sur des flux de marchandises collectés dans le cadre d'enquêtes réalisées par des fournisseurs de données ou par les opérateurs suivants : la SNCF ; les Voies Navigables de France (VNF), EUROSTAT et les Douanes (voir Encart A.9-2).

Les données disponibles dans SitraM sont fournies par « nature de transport » : national, international de type chargé ou déchargé en France, ou de transit sur le territoire français. Les marchandises sont ventilées par nature de produit selon la Nomenclature²³⁸ des Statistiques de Transport (NST), par type

²³⁷ Ce service dispose du site internet suivant : <http://www.statistiques.developpement-durable.gouv.fr>

²³⁸ La nature des marchandises est relative à leur référence dans une liste générale de produits, celle de la Nomenclature des Statistiques de Transport (NST) qui est normalisée à l'échelle internationale, elle comprend différents niveaux d'agréations des 176 types de produits élémentaires (Savy, 2007 : 54-56).

de conditionnement, par mode de transport (rail, route, voies navigables intérieures, air), par région d'origine et de destination pour le transport national français ou bien entre les régions françaises et les pays étrangers pour le transport international. Les données sont, par conséquent, plus fines lorsqu'elles sont observées depuis, dans vers ou via la France.

Encart A.9-2. Les fichiers source de SitraM

D'après le SOeS, les fichiers sources de SitraM sont :

- Le fichier TRM, l'enquête sur l'utilisation des véhicules routiers de marchandise réalisée à l'échelle nationale depuis 1975, et internationale depuis 1990 ;
- Le fichier RAIL (SNCF) de transports de marchandises par voie ferroviaire pour les wagons complets, à l'exclusion des colis individuels transportés par le SERVICE NATIONAL de Messagerie (SERNAM) ;
- Le fichier Voies Navigables Intérieures (VNF) qui porte sur les transports nationaux et internationaux de marchandises par voie navigable intérieure, en fonction des pavillons (français ou étranger) ;
- Deux fichiers issus des enquêtes TRM fournis par EUROSTAT depuis 2000 : ces fichiers ne sont pas réalisés par l'institut mais collectés à partir de sources nationales puis harmonisés. Les données portent sur l'ensemble des pays de l'Union Européenne (UE) à l'exclusion de Malte, de la Norvège et du Liechtenstein. Le premier fichier fournit par EUROSTAT porte sur les transports de marchandises inter pays : les données sont agrégées par grands groupes de produits et par pavillon ; le second fichier porte sur l'ensemble des produits transportés par pavillon, à l'échelle régionale ;
- Le fichier du commerce extérieur de la France issu des Douanes qui fournit des données sur les flux internationaux (commerciaux) de marchandises. La source est le Document Administratif Unique (DAU) avant 1988 et, depuis le 1er janvier 1993, la Déclaration d'Echanges de Biens (DEB). A noter que les données des fichiers de la DAU et de la DEB ne sont pas comparables, la méthode de mesure ayant varié. Depuis 2006, les expéditeurs ne sont plus tenus de déclarer les tonnages intra communautaires aussi les données des Douanes sur les flux internationaux de marchandises disponibles dans SitraM sont elles, issues de redressements.

L'unité de mesure des marchandises s'appuie sur un trajet isolé défini par le chargement de la marchandise dans un véhicule, au lieu d'origine, son transport et son déchargement dans un lieu de destination. Sont éventuellement mentionnés différents lieux de rupture de charges. Si les trajets individuels des marchandises sont collectés, seule une information agrégée est disponible à l'utilisateur par le biais du SitraM.

« Aussi recueille-t-on de façon précise les données touchant la nature des produits transportés (quand ils ne sont pas dans des colis de messagerie ou des conteneurs ou caisses mobiles de transport combiné, 'opaques' pour le transporteur et donc pour le statisticien), les tonnages transportés [en tonnes et en tonnes.kilomètres], les lieux de chargement et déchargement, les moyens de transport utilisés. Mais on ne dispose d'aucune information sur la taille et le poids des envois individuels, sur leur valeur, ni sur la nature du trajet complet qu'ils accomplissent (trajet unique de bout en bout ou, au contraire, acheminement dans une chaîne à maillons multiples, associant éventuellement plusieurs modes de transports) » (Savy, 2007 : 53).

D'autres sources d'information sur les flux de transport de marchandises par voie routière sont disponibles. Il s'agit, par exemple, des enquêtes rétrospectives réalisées auprès d'un échantillon de propriétaires de véhicules tirés au sort dans le fichier national des immatriculations des cartes grises.

A l'échelle internationale, les flux commerciaux sont collectés par les Douanes nationales import (ou export), grâce à la DEB renseignée par l'expéditeur lors du franchissement d'une limite internationale. La procédure de déclaration diffère selon que le bien circule à l'intérieur de l'espace Schengen ou bien entre un pays membre de l'UE et un pays non membre. La création de l'Union a, en effet, supprimé le 1^{er} janvier 1993 les contrôles douaniers aux frontières des Etats membres, selon un principe de libre circulation des biens, à quelques exceptions près, et cette suppression n'est pas sans conséquences sur les statistiques qui portent sur des flux. La procédure est formée de diverses étapes : tout d'abord, la déclaration de la marchandise sur la base d'une nomenclature particulière (biens prohibés, en libre circulation ou règlementés) permet de déterminer le niveau de formalités administratives à accomplir, notamment l'application de certaines taxes. L'identification des lieux d'origine et de destination de la marchandise est réalisée dans un second temps, celle de sa valeur monétaire dans un troisième temps. La fin de la procédure consiste en la déclaration du « régime douanier », c'est-à-dire du sens de la circulation de la marchandise (en import ou en export). Les flux internationaux issus des douanes nationales sont, en effet, toujours observés du point de vue d'un pays particulier (voir Encart A.9-3).

Ainsi, à l'échelle du périmètre de l'UE, les flux de marchandises réalisés entre les états membres sont déclarés puis enregistrés par le dispositif Intrastats© d'EUROSTAT, grâce aux différentes déclarations nationales. Le *Bulletin officiel des Douanes*, la DEB, réalisée tous les ans par le Département des statistiques et des études économiques d'EUROSTAT, décrit ces échanges de biens intra communautaires. Ce bulletin présente les données communes à toutes les déclarations et celles qui sont variables.

Encart A.9-3. La mesure des flux internationaux de marchandises par les Douanes

La mesure des flux de marchandises réalisée par les Douanes nationales consiste à renseigner un formulaire qui porte sur des articles éventuellement groupés en colis, pour lesquels sont précisés : la nature du colis (textile, par exemple) ; le pays expéditeur ; le destinataire ; le déclarant en douane (le chargeur ou la personne morale émettrice) ; le moyen de transport (par camion, par exemple) ; le groupage éventuel des colis au sein d'un même camion. Cette DEB est réalisée dès lors qu'un seuil financier de valeur de la marchandise est franchi : les flux inférieurs à 150 000 € ne sont pas soumis à déclaration. Les formalités sont aussi réduites pour les marchandises dont la valeur est comprise entre 230 000 € par an pour les flux entrants dans un pays de l'Union européenne, et 460 000 € par an pour les flux sortants. A noter également l'exonération des droits de Taxe sur la Valeur Ajoutée (TVA) à l'exportation du pays émetteur de la marchandise vers un pays membre de l'Union, c'est donc au pays récepteur de déclarer cette TVA. Le double système de seuil de valeur financière et d'exonération partielle de la TVA est supposé faciliter les échanges intracommunautaires et limiter les fraudes, notamment pour les produits à forte valeur ajoutée tels que les ordinateurs et les téléphones portables. Il conduit à un nécessaire redressement des flux de marchandises intracommunautaires pour être représentatifs de la réalité.

Toutes les marchandises qui circulent dans ou via l'UE ne donnent pas lieu à une DEB. Seules celles qui portent sur un échange entre un pays membre de l'Union et un autre pays du monde sont concernées. Les marchandises qui partent de Belgique vers l'Espagne en traversant la France ne donnent pas lieu à une DEB. En revanche, des marchandises qui partent de la France vers l'Italie, pour un client suisse, font l'objet d'une DEB. L'exemple peut être rendu plus complexe si l'on considère les flux de transit. Ainsi, une marchandise en provenance de Singapour qui entre dans l'Union européenne par le Havre (en France, flux de transit entrant), y est donc dédouanée (DAU) puis livrée à un acquéreur allemand (DEB, pour le flux sortant de France) via la Belgique (au port d'Anvers, DEB, flux de transit) et les Pays-Bas (Rotterdam, DEB, flux de transit).

D'autres dispositifs de mesure des flux internationaux commerciaux sont disponibles grâce aux statistiques de la balance des paiements et des soldes extérieurs des pays.

La balance des paiements est un décompte statistique de l'ensemble des échanges économiques et financiers réalisés pendant un an entre les *résidents* et les *non-résidents* d'un pays particulier. Il se décompose en différents comptes²³⁹. Pour la France, par exemple, le compte des transactions courantes de cette balance des paiements enregistre entre autres, les échanges de biens, de services, de revenus et de transferts courants. Les échanges de biens – qui sont ceux qui nous intéressent ici – concernent l'ensemble des opérations (de transport, logistiques, financières, ...) réalisées sur des marchandises qui franchissent la limite nationale. Les données qui concernent les soldes extérieurs sont issues de la balance des paiements. Il s'agit du transfert de marchandises depuis un pays vers un autre et qui correspond, pour le premier, à un débit et pour le second, à un crédit : les deux valeurs devant théoriquement être équilibrées afin que la balance présente un solde extérieur à zéro. Cependant, certains soldes extérieurs ne sont toutefois pas équilibrés : c'est notamment²⁴⁰ le cas du solde extérieur de la balance commerciale qui correspond aux flux de marchandises. Ce solde correspond à la différence entre les flux entrants (les importations) et les flux sortants (les exportations) d'un pays.

Enfin, il convient de mentionner l'existence de nombreuses²⁴¹ sources de données sur les flux de marchandises à l'échelle internationale. Ces bases de données sont réalisées par des organisations internationales telles que la base de données Marchandises du Fonds Monétaire International (FMI), la base de données Comptes Harmonisés sur les échanges et l'économie mondiale (CHELEM) élaborée par le Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) ou encore la base flux commerciaux de l'Organisation Mondiale du Commerce (OMC).

²³⁹ Ces différents comptes sont le Compte des transactions courantes ; le Compte de capital qui enregistre i) les transferts financiers (investissements direct étrangers (IDE), investissements des étrangers (remises), aides à l'investissement des fonds structurels européens, etc.) et ii) les acquisitions d'actifs non financiers ; et le Compte courant qui comptabilise les flux financiers (investissements directs, produits financiers dérivés,...) et les avoirs de réserve (en or, par exemple).

²⁴⁰ Les autres soldes extérieurs qui ne sont pas systématiquement équilibrés sont le solde de la balance commerciale au sens large qui concerne les flux de biens et de services ; le solde des transactions courantes qui porte sur des flux de biens, de services, de revenus et des flux financiers ; la capacité de financement de la nation obtenue par la somme du solde des transactions courantes et du solde du compte de capital ; et le solde des flux financiers qui concerne les flux de capitaux, hormis les avoirs de réserves.

²⁴¹ Outre la base de données (BD) Marchandises de la Direction des Statistiques sur le Commerce du Fonds Monétaire International (FMI), on peut citer les BD Marchandises de l'Organisation de Coopération et de Développement Economiques (OCDE), de la Conférence des Nations Unies sur le commerce et le développement (CNUCED) et de la Banque Mondiale ; la BD des trafics portuaires commercialisée par *Containerisation Internationale* ou encore celle de l'ISL ; la BD de l'Organisation Mondiale du Commerce (OMC) sur les échanges par catégories de produits ou encore la BD sur les trafics portuaires à l'échelle européenne de la Commission des Ports Flamands (VHC Ports).

Annexe A. 2. La construction d'une matrice de voisinage spatial (*k* plus proches voisins)

Cette annexe porte sur la construction d'une matrice de voisinage devant permettre de réduire une matrice de flux, dans l'objectif de les cartographier en fonction d'un critère d'éloignement discret. Elle intéresse le chapitre 7.

Cette matrice de voisinage concerne le *plus proche voisin*, elle est élaborée à partir d'un fond de carte ponctuel et non surfacique comme dans le chapitre 7. L'objectif est de montrer la variation du *graphe de voisinage* obtenu en fonction du type d'implantation spatiale considéré : ponctuel ou linéaire (voir section 7.2 du Chapitre 7).

La mesure du voisinage selon le *plus proche voisin* associe donc ici la localisation (x,y) des barycentres des unités spatiales, avec une mesure de proximité géographique continue.

A.2.1 Présentation de l'approche

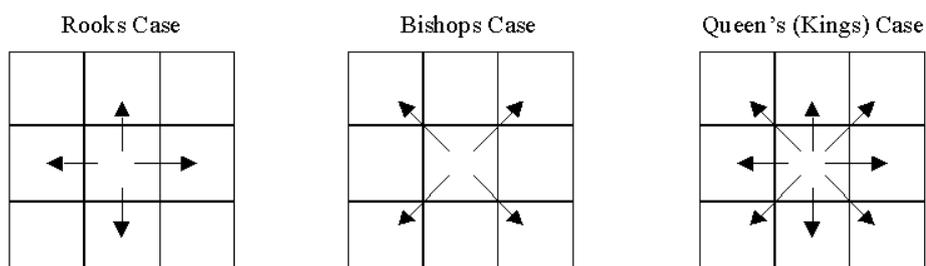
L'approche mobilise une *matrice des k-contiguïtés*, parfois qualifiée dans la littérature de matrice de *poids spatiaux*. Cette matrice, généralement utilisée dans les analyses de régression spatiale, sert à déterminer l'importance et à la portée spatiale d'une mesure d'auto corrélation spatiale des données, c'est-à-dire de l'extension dans l'espace du degré de ressemblance d'une unité spatiale donnée à ses voisins. On définit généralement à partir de cette *matrice des k-contiguïtés*, une *matrice de distances* pouvant être discrète ou continue (par exemple de $1/D_{ij}$ afin de considérer la décroissance de l'effet de la distance), ou une *matrice de pondération spatiale* (W_{ij}), qui définit la part en pourcentage du poids du lien dans la somme marginale en ligne : $W_{ij}=W_{ij}/O_i$, ou encore le total général. Cette matrice servira à tester l'auto corrélation spatiale.

Fondée sur la notion de *voisinage*, la mesure *des k-contiguïtés* est donc composée de deux étapes. Elle nécessite de définir le critère de voisinage à partir duquel deux unités spatiales seront considérées comme voisines ou proches. Le résultat est stocké dans une matrice de voisinage qui prend la forme d'une matrice d'adjacence : les couples de voisins prennent la valeur '1' sinon 0. La seconde étape vise à définir le poids attribué à chacun des liens de cette matrice de voisinage, compte tenu de leur proximité au regard de la thématique du flux, correspondant à la *matrice de poids spatiaux*.

Seule la première étape est nous intéresse ici dans cette approche cartographique de sélection de valeurs de flux intégrant le voisinage des lieux.

Le *voisinage* d'un lieu peut en effet être défini de deux manières : en termes de *contiguïté* des zones selon qu'elles partagent une frontière commune ($C_{ij}=1$) ou un point. Dans ce cas, elle tient également compte de la distance à un nombre *k* de *plus proche voisin*. Le premier cas correspond ainsi, sous certaines conditions, à la matrice (C_{ij}) dont il a été question dans la section 7.2.1. Il s'agit de l'exemple de droite représenté sur la Figure A.2-30., qui illustre le déplacement de la Reine sur un échiquier.

Figure A.2-30. Méthodes d'estimation du voisinage spatial fondé sur la contiguïté des zones



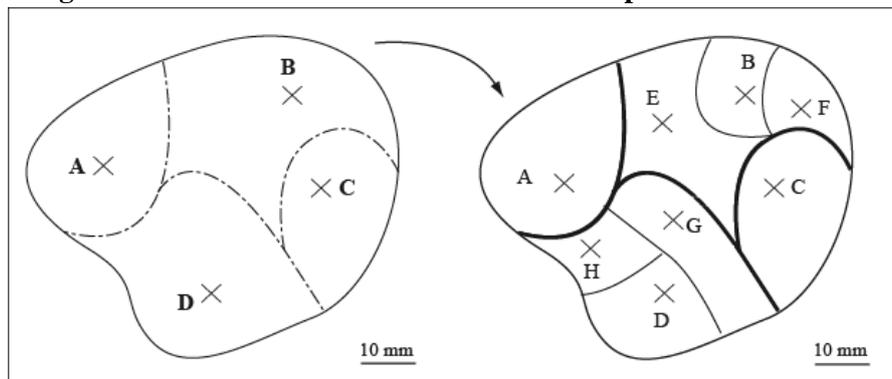
La Figure A.2-30 présente les trois possibilités théoriques de sélection du plus proche voisin (j), en fonction de l'orientation de (i). Elle est souvent envisagée, dans la littérature, en référence aux déplacements de la Tour, du Fou et de la Reine sur un échiquier.

La différence par rapport à l'approche présentée dans la section 7.2.1 réside dans le fait que la recherche de *plus proche(s) voisin(s)* considère une implantation ponctuelle de la zone d'étude — la matrice des points de coordonnées (x, y) — et non surfacique comme dans le cas de $(C_{ij}=1)$ et une mesure d'éloignement continu. Le graphe de voisinage spatial correspond, en effet, à une matrice où la valeur d'une cellule (i,j) est soit celle d'un nombre k de (j) voisins géographiques considérés dans un périmètre donné au regard d'une mesure de distance (continue, généralement (D_{ij}) ou discrète, correspondant à une distance géodésique, par exemple (CK_{ij}) telle que présentée dans la section 7.2.2) ; le choix de la mesure d'éloignement étant spécifique à la thématique. On peut, par exemple, recourir à une distance sociale ou géographique pour déterminer le nombre (k) de voisins, en considérant un ordre où une portée de distance : par exemple le nombre (k) de voisins situés dans un rayonnage de α km ou inférieur au premier quartile $(D_{ij} < Q1)$.

Ainsi, de la même façon que la matrice des k voisins avec $(k=1)$ ne correspondait à $(C_{ij}=1)$ (1 frontière) que dans le cas de la Reine (voir Figure A.2-30), on observe des différences importantes lorsque l'on considère des portées discrètes plus importantes $(k=1, 2, 3, \dots, n)$. Pour illustrer ces différences, nous rappelons la formation de la matrice d'exemple local à huit lieux.

La Figure A.2-31 illustre la construction de ces deux matrices dont l'ordre de voisinage est finalement d'ordre 1, la valeur de k variant fondamentalement entre les deux matrices.

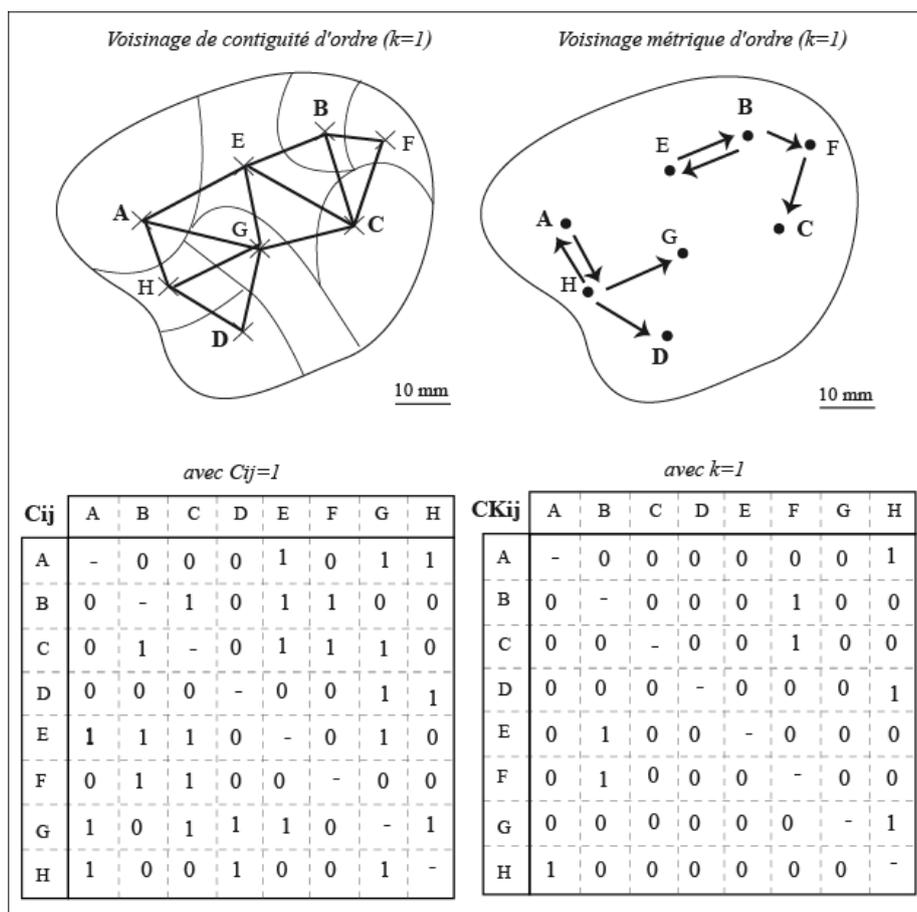
Figure A.2-31. Présentation de la matrice exemple local à huit lieux



Pour bien comprendre les différences entre les matrices $(C_{ij} \neq CK_{ij}^{k=1})$ qui sont finalement toutes deux d'ordre 1 ($k=1$), nous allons supposer que les 4 points de notre matrice d'exemple local correspondent, en réalité, aux chefs-lieux d'un découpage administratif territorial de niveau régional que nous mobiliserons dans la prochaine section. Cela nous permet dans l'immédiat de considérer une zone d'étude au niveau local qui compte huit lieux. C'est cette modification de l'exemple local qui est présentée sur la Figure 7-12.

Sur la Figure A.2-32, l'image de gauche correspond à la matrice (C_{ij}) alors que celle de droite à (CK_{ij}) avec $(k=1)$, présentée dans le chapitre 7. La première s'appuie sur une mesure binaire (l'existence d'une frontière) mobilisant ainsi un maillage spatial discret ; la seconde requiert une mesure de distance, que nous avons supposé être exprimée en (kilo)mètres (mais plusieurs mesures sont envisageables), pour identifier l'ordre de voisinage considéré. Compte tenu de ces différences, nous proposons d'utiliser les expressions respectives de *voisinage de contiguïté* et de *voisinage métrique* (dans les deux cas : *d'ordre $k=1$*) pour distinguer littéralement ces deux matrices.

Figure A.2-32. Illustration du voisinage de contiguïté versus voisinage métrique. Comparaison des méthodes d'estimation des $k=1$ (plus proches) voisins



Trois différences majeures sont notables sur le graphe de voisinage issu de ces deux matrices (voir Figure A.2-3), lesquelles expliquent par ailleurs la forte dissemblance de leurs motifs spatiaux et de leur sémiologie graphique. La première tient au nombre de liens considérés pour chacun des lieux ; la seconde à leur caractère symétrique ou non par rapport à la diagonale principale et la troisième, à l'élimination de certains lieux de l'analyse.

La matrice (CK_{ij}) avec ($k=1$), située sur la partie droite de la Figure A.2-32 ne retient qu'un seul voisin par lieu : le plus proche en distance ; il s'agit du *premier plus proche voisin* qui n'est représenté que par un lien, pour chacun des lieux. A l'inverse, pour la matrice (C_{ij}), un même lieu peut avoir plusieurs voisins avec lesquels il partage une frontière, tout dépend de la morphologie des zones d'étude, *a fortiori* du type de maillage utilisé pour l'analyse : l'utilisation d'un maillage régulier ou théorique, tel un pavage de Voronoï, ne conduira pas au même résultat que l'utilisation d'un maillage hétérogène.

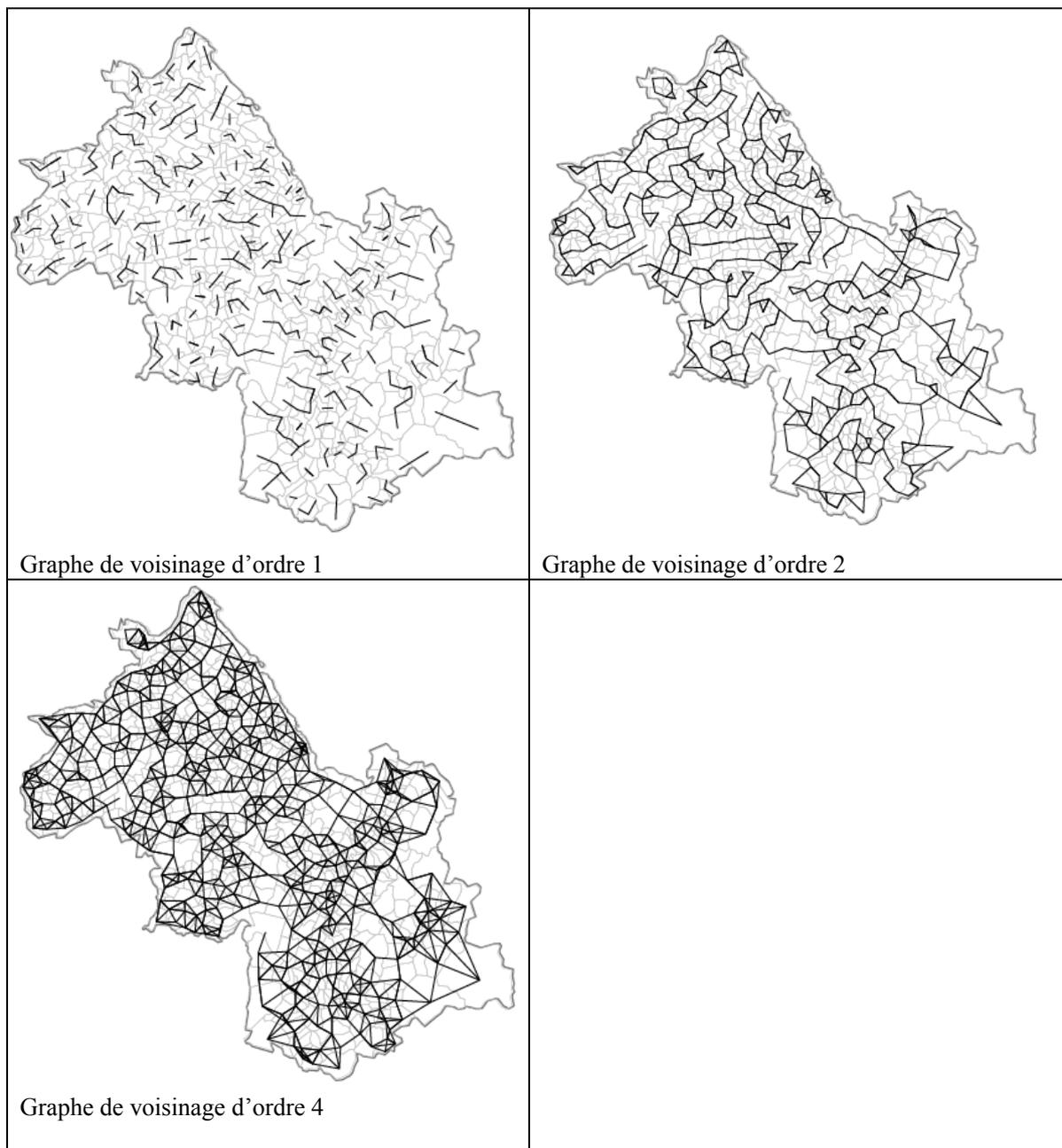
La deuxième différence et non la moindre entre ces deux mesures de proximité de voisinage d'ordre 1, réside dans la non symétrie de la matrice (CK_{ij}), la matrice (C_{ij}) étant quant à elle symétrique. On observe, en effet, sur la Figure A.2-32, que le plus proche voisin de (E) est (B) tandis que le plus proche voisin de (B) est (F). De la même façon, certains lieux, tels que (C) ou (G) par exemple, ne sont le plus proche voisin de personne, même s'il a nécessairement un voisin le plus proche. Cela s'explique par le caractère asymétrique de cette matrice.

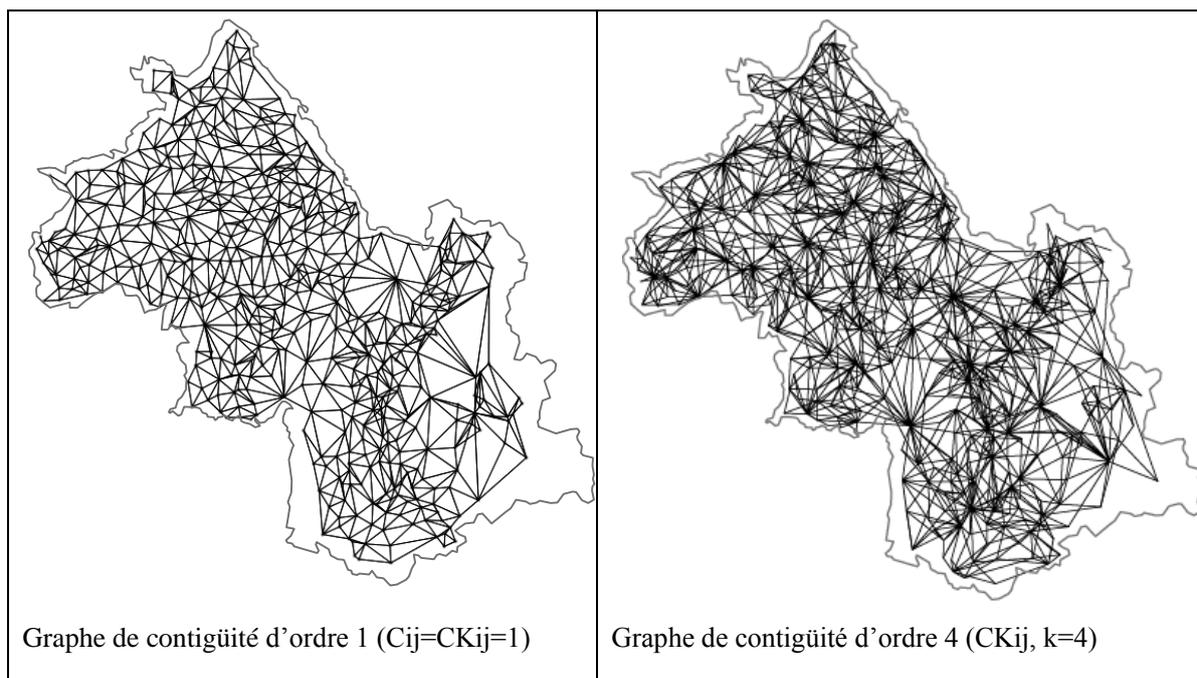
A.2.1 Comparaison des différents graphes de voisinage de l'Isère

La Figure A.2-33 applique la procédure de la Figure A.2-3 au fond de carte des communes de l'Isère, pour $(k=1)$, $(k=2)$ et $(k=4)$, elle présente les graphes de voisinage d'ordres 1, 2 et 4.

La matrice (CK_{ij}) avec $(k=1)$ étant *asymétrique*, sa représentation graphique nécessite le recours à la flèche afin d'identifier sans confusion le lieu depuis lequel le plus proche voisin est considéré. C'est la raison pour laquelle la sémiologie des motifs de la spatialisation de ces deux *graphes de voisinage* diffère.

Figure A.2-33. Graphes de voisinage et de contiguités d'ordres $k=(1)$, $k=(2)$ et $k=(4)$ du fond de carte de l'Isère





D'autres différences peuvent être mentionnées, parmi lesquelles le choix de la direction en degrés de recherche du plus proche voisin ; celui d'un maillage théorique²⁴² permettant d'identifier le (ou les) voisin.s concerné.s ou la métrique utilisée. Dans le cas de (C_{ij}) , nous avons vu que la métrique est binaire et mobilisée seule, alors que dans le cas de (CK_{ij}) avec $(k=1)$, la mesure est binaire, ordinale et toujours associée à une métrique d'éloignement (CK_{ij}, D_{ij}).

L'utilisation de la matrice (CK_{ij}) ouvre aussi la voie vers la construction de matrices des deux plus proches voisins jusqu'à $(k=n)$ plus proches voisins.

Le programme pour Rstudio© permettant la construction des matrices présentées est disponible en annexe, ainsi que sur le site du groupe de travail ElementR.

²⁴² Les travaux menés en écologie, sur des implantations ponctuelles, mobilisent des maillages et différentes conditions. Par exemple : deux points seront considérés comme voisins s'ils partagent une frontière dans le pavage de Delaunay et si aucun autre point ne se trouve à l'intérieur du diamètre qu'ils définissent (Gabriel et Sokal, 1969), cités par <http://pbil.univ-lyon1.fr/R/pdf/ter4.pdf> (page 5)

Annexe A. 3. La Transformation d'une matrice (Fij) en une matrice ordinale

Cette annexe permettant de transformer une matrice binaire en une matrice de rangs intéresse le chapitre 9.

La méthode de J. D. Nyusten et M. F. Dacey (1961) s'inscrit dans le cadre de la théorie des graphes, de même que dans celui des statistiques non paramétriques, qui ne postulent pas la conformité de la série à une loi connue. Le recours à ce dernier cadre théorique autorise la sélection des valeurs de flux à cartographier, en fonction de leur rang, c'est-à-dire de leur position (1, 2, ..., n) dans la distribution des valeurs. Il convient de ne pas confondre l'analyse des rangs d'une matrice, avec le rang d'une matrice. Nous rappelons que le rang d'une matrice (Fij), noté $r(Fij)$ est le nombre (r) de ses (N) vecteurs lignes et de colonnes, soit de ses (N) lieux d'origine ou de destination. Une matrice dite *carrée* de rang 3 possède trois lignes et trois colonnes.

Cette approche selon les rangs à l'avantage d'être facile à mettre en œuvre, dès lors que les données ont été triées. Dans une analyse des rangs, les valeurs de la série sont généralement triées selon un ordre décroissant, puis remplacées par leur rang dans la série. Cette transformation des flux en valeurs ordinales est, en pratique, une forme de permutation de la matrice et il en existe $n!$ différentes. Cependant, toutes les permutations ne sont pas pertinentes dans le cadre d'analyse cartographique des flux, car elles conduisent dans bien des cas à la perte de la composante de localisation des lieux. Aussi allons-nous restreindre le champ d'application de cette méthode pour ne pas supprimer la composante spatiale.

La Figure A.3-34 présente la procédure générale de transformation des valeurs de flux en rangs, elle est appliquée à notre matrice exemple local.

Figure A.3-34. Illustration de la transformation des valeurs de flux (Fij) en k rangs

<i>Valeurs disponibles</i>						<i>Rangs</i>					
Fij	A	B	C	D	Oi	Fij	A	B	C	D	Oi
A	-	8	5	1	14	A	-	2	4	10 _{ex}	1
B	6	-	3	2	11	B	3	-	6 _{ex}	8 _{ex}	3
C	10	4	-	1	15	C	1	5	-	10 _{ex}	2
D	1	2	3	-	6	D	10 _{ex}	8 _{ex}	6 _{ex}	-	4
Dj	17	14	11	4	46	Dj	1	2	3	4	

2 : ordre du rang
2_{ex} : ordre du rang identique pour 2 entités et plus

La matrice des rangs située sur la partie droite de cette figure correspond à un ordonnancement global des valeurs de la matrice des données initiales (Fij). La valeur de flux (C, A)=10 étant la plus importante à l'échelle globale, elle prend le rang 1 ; la valeur de (A, B)=8 étant la seconde valeur la plus importante, elle prend le rang 2 et ainsi de suite.

L'analyse des rangs réalisée dans un objectif de cartographie des flux consiste, en effet, à trier les valeurs de flux, à leur affecter un rang puis à représenter les rangs des flux les plus importants. Cette représentation conduit à une hiérarchisation des lieux en fonction de leur position dans le système global d'échanges ainsi ordonné. Cependant, pour être efficace, l'ordonnement des valeurs de flux ne doit pas être réalisé à ce niveau global, comme dans le cas de la Figure A.3-34 il conduit à un résultat similaire à la sélection des n plus gros flux présentée la section 9.1 du chapitre 9. L'application la plus efficace de la sélection des flux selon leur rang, doit donc être locale soit réalisée du point de vue des lieux d'origine ou de destination.

Cette possibilité est présentée sur la Figure A.3-35. : les valeurs de flux émises par chacun des (i) lieux sont ainsi ordonnées localement, pour chacun des lieux.

Figure A.3-35. Illustration de la transformation des flux (F_{ij}) depuis l'origine (i) en k rangs

<i>Valeurs disponibles</i>					
F_{ij}	A	B	C	D	O_i
A	-	8	5	1	14
B	6	-	3	2	11
C	10	4	-	1	15
D	1	2	3	-	6
D_j	17	14	11	4	46

<i>Rangs</i>					
F_{ij}	A	B	C	D	O_i
A	-	1	2	3	1
B	1	-	2	3	3
C	1	2	-	3	2
D	3	2	1	-	4
D_j	1	2	3	4	

2 : ordre du rang des flux depuis (i)

Du point de vue du lieu d'origine (A), la valeur de flux (A, B)=8 étant la plus importante, elle prend le rang 1 (voir partie droite de la figure) ; la valeur (A, C)=5 ou la seconde valeur la plus importante prend le rang 2. Quant à la valeur de flux (A, D), elle prend le rang 3 et ainsi de suite pour les autres lieux.

De cette manière, l'application du critère de sélection ($k \geq 1$), avec ($k=1$) conduira à ne retenir que les flux de rang 1 (donc les *flux majeurs*) pour l'analyse des *flux dominants* selon méthode (Nyusten-Dacey).

PROGRAMMES POUR RStudio ANNEXES

Contribution à la cartographie d'une matrice de flux

Françoise Bahoken
octobre 2015

PROGRAMMES POUR RSTUDIO ANNEXES

Ce document regroupe des procédures commentées intéressant les chapitres 3, 7, 8 et 9.

Sommaire :

1. Procédure de transformation d'une matrice de flux fermée non carrée en une matrice carrée.....	2
2. Procédure de cartographie d'une matrice de flux (approche matricielle).....	6
3. Procédure de sélection des flux à cartographier, selon un critère spatial (de voisinage binaire, ordinal ou de distance parcourue).....	12
4. Procédure de sélection des flux à cartographier, selon un critère global.....	22
5. Procédure de sélection des flux à cartographier, selon un critère local (« flux dominants »).....	28
6. Procédure de cartographie directe d'une variable de l'environnement global de Rstudio©.....	35
7. Références.....	38
8. Plan détaillé	39

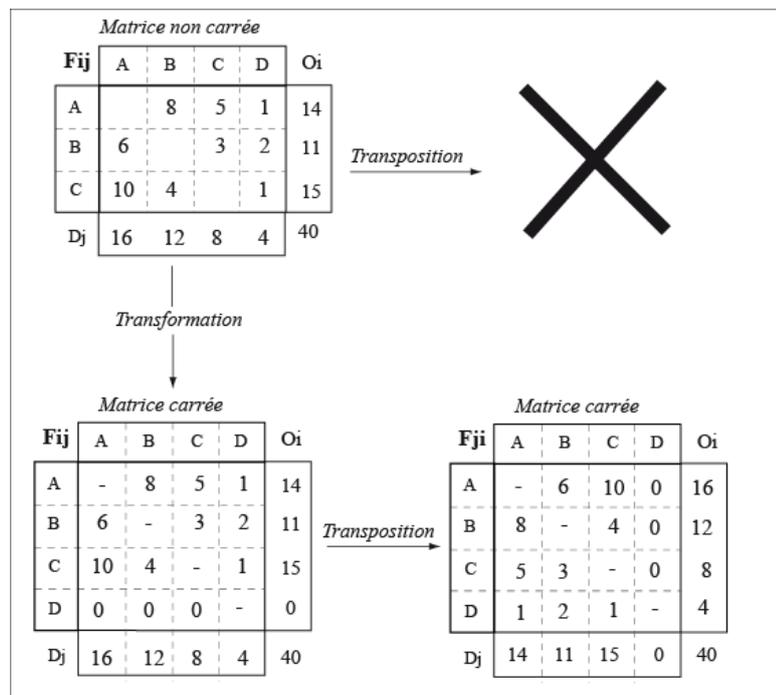
1. Procédure de transformation d'une matrice de flux fermée non carrée en une matrice carrée

Cette annexe décrit une procédure permettant de transformer une matrice non carrée ($n*m$), formée de n lignes et de m colonnes d'effectifs différents, en une matrice carrée où le nombre de lignes est égal au nombre de colonnes ($n*n$). On utilise pour cela une liste (n) d'identifiants contenant l'ensemble des lieux de la zone d'étude. Le cas d'illustration est la matrice de flux domicile-travail de l'Isère qui n'est pas « carrée ».

1.1 Justification

Lorsqu'une matrice n'est pas carrée, que le nombre de lignes et de colonnes diffère, il n'est pas possible de la décomposer selon la procédure présentée dans le Chapitre 3, en utilisant sa propriété générale de permutation des valeurs au regard de la diagonale principale. La transposition d'une matrice non carrée n'ayant en fait pas de sens (voir Figure R-1-1).

Figure R-1-1. Transformation d'une matrice de flux non carrée en une matrice carrée



Pour pouvoir permuter une matrice non carrée, il faut donc la transformer, la rendre carrée. Cette opération est en pratique délicate, car elle consiste à augmenter la matrice en introduisant une (ou des) ligne(s) et/ou une (ou des) en colonne(s) vides. Autrement dit de générer des couples de lieux dont les valeurs n'ont pas été observées et qui seront donc portées à 0. Toutefois, l'information de flux déjà disponible ne sera ni modifiée, ni transformée ni altérée, l'opération n'affectant que le format du tableau.

Pour l'illustrer, nous allons prendre l'exemple de la matrice de l'Isère, formée de 532*498 individus en sortie directe du fichier national MOBPRO.

La procédure consiste à générer une matrice carrée, à partir de la liste des identifiants de l'ensemble des unités spatiales de la zone d'étude. Pour l'Isère, il s'agit de la liste des 533 communes. La matrice d'origine (532*498) sera donc transformée en matrice (533*533).

La première étape consiste à générer une matrice carrée à valeurs nulles ($F_{ij}=0$) de rang 533. Cette matrice sera ensuite renseignée en utilisant les données de la matrice (532*498), en contrôlant simultanément les codes des lieux d'origine et de destination par la fonction *merge*. Afin de ne pas introduire d'erreurs, nous avons généré un code identifiant unique pour chacun des couples de lieux, issu de la concaténation des codes des lieux d'origine et de destination.

1.2 Préparation

1.2.1 Connexion au répertoire de travail et chargement des packages requis.

```
#purge environnement de travail
rm(list=ls())
setwd("D:/FRANCOISE/R/Isere/transfo/")

# declaration des packages requis
library(reshape2)
library(rgdal)
```

1.2.2 Chargement de la liste des identifiants

On charge une liste des codes INSEE des communes

```
tab <- read.csv2("CODE_Isere.csv", header=TRUE, sep=";", stringsAsFactors=FALSE, encoding="UTF-8", dec=",", check.names=FALSE)
str(tab)
## 'data.frame': 533 obs. of 1 variable:
## $ CODE: int 38001 38002 38003 38004 38005 38006 38008 38009 38010 38011 ...
```

1.3 Construction d'une matrice vide et transformation de la matrice non carrée (format matriciel)

1.3.1 Construction d'une matrice vide

Construction d'une matrice (*mat*) à 0, de dimension ($n*n$) où n est le nombre de lignes correspondant au nombre d'identifiants

```
nbi<-dim(tab)[1]
mat <- matrix(data = c(0), nrow = nbi, ncol =
nbi, dimnames=list(c(as.matrix(tab)),c(as.matrix(tab))))
class(mat)
## [1] "matrix"
```

1.3.2 Transformation de cette matrice au format liste

Transformation de cette matrice en une liste (3 colonnes) et création d'un nouvel identifiant unique pour chaque couple de lieu (*ID_link*), issu de la concaténation des codes des lieux d'origine et de destination

```
tabflow <-melt(mat)
names(tabflow) = c("j", "i", "Fij")

tabflow$ID_link<-paste(tabflow$i,tabflow$j, sep = "-")

head(tabflow)
```

```
##      j      i Fij      ID_link
## 1 38001 38001  0 38001-38001
## 2 38002 38001  0 38001-38002
## 3 38003 38001  0 38001-38003
## 4 38004 38001  0 38001-38004
## 5 38005 38001  0 38001-38005
## 6 38006 38001  0 38001-38006
```

1.3.3 Chargement de la matrice des données de flux observées non carrée (au format liste)

Chargement de la liste des valeurs de flux observées (data), dont on sait qu'elle n'est pas carrée.

Création d'un nouvel identifiant unique pour chaque couple de lieu (ID_link), comme précédemment.

Le lieu d'origine =COMMUNE et le lieu de destination=DCLT, la variable COUNT correspond à la valeur du flux.

```
data <- read.csv2("source.csv",header=TRUE,sep=";",stringsAsFactors=FALSE,encoding="UTF-8",dec=".",check.names=FALSE)
data$ID_link<-paste(data$COMMUNE,data$DCLT,sep = "-")
```

```
head(data)
##  COMMUNE DCLT  COUNT  ID_link
## 1  38001 38001 527.7349 38001-38001
## 2  38001 38012  15.3317 38001-38012
## 3  38001 38013   7.6658 38001-38013
## 4  38001 38022  30.6634 38001-38022
## 5  38001 38028   3.8329 38001-38028
## 6  38001 38029  11.4988 38001-38029
```

1.3.4 Transformation des données de flux observées en matrice carrée

Réalisation d'une jointure entre la matrice vide (tabflow) et la matrice de données observées (data), en utilisant comme clé de jointure : ID_link, puis construction de la matrice (Fij), dans un format liste.

```
tabflow= data.frame(tabflow, data[match(tabflow["ID_link"], data["ID_link"]),2:3])
Fij<-data.frame(tabflow$i,tabflow$j,tabflow$COUNT)
names(Fij) = c("i", "j", "COUNT")
head(Fij)
##      i      j  COUNT
## 1 38001 38001 527.7349
## 2 38001 38002      NA
## 3 38001 38003      NA
## 4 38001 38004      NA
## 5 38001 38005      NA
## 6 38001 38006      NA
```

Transformation de la matrice (Fij) du format liste au format matriciel, vérification du format et annulation des valeurs de la diagonale

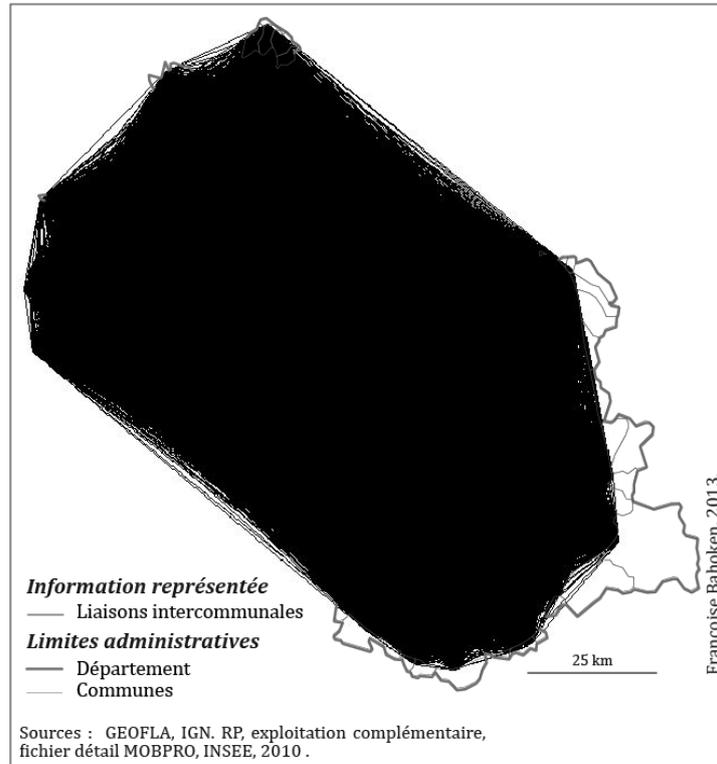
```
matflow<-acast(Fij,i~j)
## Using COUNT as value column: use value.var to override.
dim(matflow)
## [1] 533 533
diag(matflow)<-0
```

Les cases vides de cette matrice étant renseignées par NA, on les met à zéro.

```
for (i in 1:nrow(matflow))
  for (j in 1:ncol(matflow))
    {if (is.na.data.frame(matflow[i,j])==TRUE) {matflow[i,j]<-0}}
```

Cartographie des liaisons théoriquement disponibles de la matrice des navettes de l'Isère (533*533), soit 284 089 figurés, en utilisant le programme présenté dans l'annexe R-2 – Cartographie (voir Figure R-5-2).

Figure R. 1-2. La géovisualisation des liaisons théoriques de la matrice de l'Isère (2010)



Export de la matflow en (Fij), pour traitements ultérieurs

```
#write.table(Fij,file="./matflow.csv",sep=";",eol="\n")
```

2. Procédure de cartographie d'une matrice de flux (approche matricielle)

Cette annexe est transversale aux chapitres 3, 7, 8 et 9. Elle permet de cartographier n'importe laquelle des matrices de liens ou de flux, réduite ou non. La préparation des données est supposée avoir été réalisée en amont de cette étape de représentation. Une version réticulaire, mobilisant un graphe est disponible sur le site du groupe ElementR.

Le cas d'application est la matrice de l'Isère.

2.1 Préparation

Connexion au répertoire de travail et chargement des packages requis.

```
#purge environnement
rm(list=ls())

#déclaration des packages requis
library(reshape2)
library(gdata)
library(rCarto)
library(maptools)
library(RColorBrewer)
library(classInt)

#connexion à L'espace de travail
setwd("D:/FRANCOISE/R/Isere")
```

2.1.1 Chargement des données numériques

Chargement de la matrice à analyser

```
# Matrice Isère - flux domicile-travail (en nombre d'actifs de 15 ans et plus)
tabflow <- read.csv2("./liste_isere.csv",header=T,sep=";",stringsAsFactors=T,encoding="UTF-8",dec=".", check.names=F)
head(tabflow)
##      i      j Fij
## 1 38001 38001  0
## 2 38002 38001  0
## 3 38003 38001  0
## 4 38004 38001  0
## 5 38005 38001  0
## 6 38006 38001  0
```

Pour cartographier les flux, il est nécessaire de charger une matrice sous la forme de liste (dataframe) comme ci-dessus, sinon, il convient de transformer la matrice, en utilisant la fonction `melt` (matrice → liste). A noter que la transformation d'une liste en matrice (liste → matrice) s'effectue avec la fonction `dcast`.

Transformation (éventuelle) de la matrice en liste, si elle ne l'est pas.

```
#Fij <- as.matrix(tab[, -1])
#row.names(Fij) <- tab[, 1]
#class(Fij)

#matflow <- Fij
#tabflow <- melt(matflow)
#names(tabflow) = c("i", "j", "Fij")
#head(tabflow)
```

2.1.2 Préparation des matrices décomposées (F_{ij}^+), (F_{ij}^-) – forme matricielle

Ces différentes matrices sont issues de la décomposition éventuelle de (F_{ij}) en : (F_{ij}), (F_{ji}), (F_{ij}^+), (F_{ij}^-), comme indiqué dans le Chapitre 3. Deux possibilités sont présentées, selon que la matrice est disponible sous la forme d'une matrice (matflow) ou bien d'une liste (tabflow).

2.1.2.1 Décomposition de (F_{ij}), dans le format matriciel

F_{ji} , $FS_{ij}=(F_{ij}^+)$ et $FD_{ij}=(F_{ij}^-)$

```
#Fji<-t(Fij)
#FSij <-Fij+Fji
#FDij <-Fij-Fji
```

2.1.2.2 Décomposition de (F_{ij}), dans le format liste

On note $tabflow=(F_{ij})$; $(F_{ij}^+)=FS_{ij}$ et $(F_{ij}^-)=FD_{ij}$.

```
#tab <- tabflow
#i<-tab$i
#j<-tab$j
#Fij<-tab$TOT
#head(tab)
##      i      j  Fij
## 1 38001 38012  15
## 2 38001 38013   7
## 3 38001 38022  30
## 4 38001 38028   3
## 5 38001 38029  11

#Fij_1<-data.frame(i,j,Fij)
#Fji_2<-data.frame(i,j,Fji)
names(Fji)<-c("i","j","Fji")

#Fij<-merge(Fij_1, Fji_2,by=c("i","j"),aLL.X=TRUE,aLL.Y=TRUE)
#Fij$FSij<- Fij$Fij+ Fij$Fji
#Fij$FDij<-Fij$Fji-Fij$Fij

#head(Fij)
##      i      j  Fij  Fji  FSij
## 1 38001 38012  15  15   30
## 2 38001 38013   7   7   14
## 3 38001 38022  30  30   60
## 4 38001 38028   3   3    6
## 5 38001 38029  11  11   22
## 6 38001 38053  45  45   90
```

On fait le choix de poursuivre l'analyse avec la matrice $(F_{ij})=tabflow$, présentée sous la forme d'une liste.

2.1.3 Chargement des données cartographiques (zones et/ou points)

Import et affichage du fond de carte (fdc) au format shapefile (*.shp) des zones (ici pour l'habillage) et éventuellement, des points de coordonnées (pt) s'ils sont disponibles, sinon on les calculera.

```
fdc <- readShapeSpatial("./Isere_com.shp")
#pt<-readShapeSpatial("./Isere_coord.shp") # chargement d'un fichier de points de
coordonnées (x,y)
```

Si le fichier de points de coordonnées n'est pas disponible, on le prépare comme ci-dessous.

Extraction des centroïdes des zones et création d'un fichier de Points (pt) comprenant leur code identifiant (CODE), leurs coordonnées en X et en Y.

```
pt <- cbind(fdc@data[, "INSEE_COM"], as.data.frame(coordinates(fdc)))
colnames(pt) <- c("CODE", "X", "Y")
head(pt)
##      CODE      X      Y
## 0 38226 917543.1 6416638
## 1 38206 933896.6 6464035
## 2 38301 907669.5 6413449
## 3 38448 840772.2 6484481
## 4 38287 893903.3 6483937
## 5 38153 904873.1 6458243
```

Réalisation d'une jointure attributaire entre la liste de flux (tabflow) et le fichier de points (pt)

```
tabflow= data.frame(tabflow, pt[match(tabflow[, "i"], pt[, "CODE"]), 2:3])
tabflow = data.frame(tabflow, pt[match(tabflow[, "j"], pt[, "CODE"]), 2:3])
colnames(tabflow) <- c("i", "j", "Fij", "X1", "Y1", "X2", "Y2")
head(tabflow)
##      i      j Fij      X1      Y1      X2      Y2
## 253 38001 38001  0 902347.2 6496924 902347.2 6496924
## 63 38002 38001  0 934589.3 6466585 902347.2 6496924
## 260 38003 38001  0 845671.2 6472517 902347.2 6496924
## 381 38004 38001  0 892928.9 6460737 902347.2 6496924
## 51 38005 38001  0 937063.9 6456949 902347.2 6496924
## 47 38006 38001  0 943664.5 6480414 902347.2 6496924
```

2.2 Cartographie des liens (Lij)

Affichage de l'ensemble des liens théoriques (Lij) – Attention, si la matrice est dense, cette figure nécessite des ressources machine conséquentes pour pouvoir être affichée.

On indique tout d'abord que l'on souhaite superposer différents objets graphiques sur la même figure, en l'occurrence un fond de carte (fdc) et des lignes (segments) ou des flèches (arrows).

```
par(mfrow=c(1,1))
```

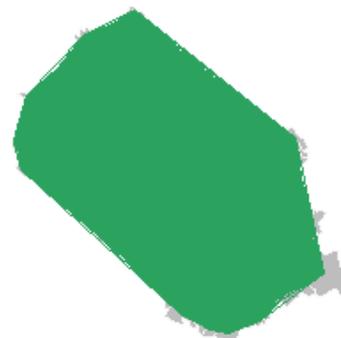
Puis on représente ces deux objets graphiques

```
plot(fdc, col="#bdbdbd", border="#f7f7f7", lwd=0.5)
segments(tabflow$X1, tabflow$Y1, tabflow$X2, tabflow$Y2, col="#2ca25f", lwd=tabflow$size)
title(main="Carte de Flux de l'Isere ", sub="(c) Françoise Bahoken R 2015", cex.sub=0.9)
```

Il est important de noter que la représentation graphique que l'on obtient correspond aux liens théoriques. En pratique, la matrice de flux correspondante n'est pas complète (certaines cases ne sont pas renseignées).

Pour visualiser les valeurs de flux de cette matrice (l'information de flux réellement disponible), quatre choix de représentation doivent ensuite être effectués à ce stade :

(1) Choix des liens valués (des cases de la matrice renseignées) ;



- (2) Choix du seuil d'affichage des valeurs des flux ($F_{ij} > \alpha$), dans le cas où la matrice n'a pas été préalablement filtrée ou réduite ;
- (3) Choix de la taille maximale de la *largeur* du figuré ;
- (4) Choix du degré d'opacité de la teinte (application éventuelle d'une *Transparence* aux figurés).

2.3 Cartographie de flux (supérieurs à un critère ($F_{ij} > \alpha$))

On présente ici une manière de sélectionner les flux à représenter à partir d'un unique critère global (Chapitre 8).

2.3.1 Choix du critère de sélection des valeurs de flux

La cartographie de flux consiste, ici, à appliquer un critère statistique en deçà duquel les flux ne seront pas représentés. Pour choisir ce critère, on effectue un résumé statistique, pour choisir par exemple, la valeur du flux moyen.

```
summary(tabflow$Fij)
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.000  0.000  0.000  1.036  0.000 4002.000
```

Il est également possible de déterminer une valeur quelconque, qu'on définit empiriquement. On décide ici de ne retenir que les flux supérieurs à 100 (ils correspondent au choix réalisé par l'INSEE de mise à disposition des données sur son site internet).

```
nb<-100 # seuil de flux
tabflow_seuil<-tabflow[tabflow$Fij>=nb,]
```

ou

Choix du seuil version (2)

```
#summary(tabflow$Fij)
#seuil <-mean(tabflow$Fij) #on choisit un seuil
#tabflow_seuil <-tabflow[tabflow$Fij>=seuil,]
```

2.3.2 Choix du paramètre de taille graphique des flux (largeur du figuré)

On décide que la largeur du plus gros flux n'excèdera pas 20 points. Deux versions sont proposées :

- (1) soit on ne dimensionne que la taille maximale du plus gros flux ;
- (2) soit on paramètre les tailles minimales et maximales des plus gros et plus petits figurés.

Paramétrage de la *largeur* maximale

```
maxsize<-20 #épaisseur du plus gros lien
#tabflow_seuil$size<-(tabflow_seuil$Fij/max(tabflow_seuil$Fij))*maxsize
```

Paramétrage des *largeurs* minimale (*minsize*) et maximale (*maxsize*) des figurés

```
#Fij_sel<-tabflow_seuil

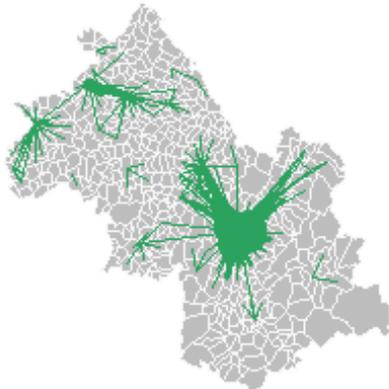
#maxsize<-20
#minsize<-0
#size$size<-minsize+(maxsize-minsize)*(Fij_sel$Fij-min(Fij_sel$Fij))/(max(Fij_sel$Fij)-min(Fij_sel$Fij))
```

2.3.3 Tracé des flux conformément aux deux critères (graphiques et thématiques)

Tracé des flux conformément aux deux critères (graphiques et thématiques)

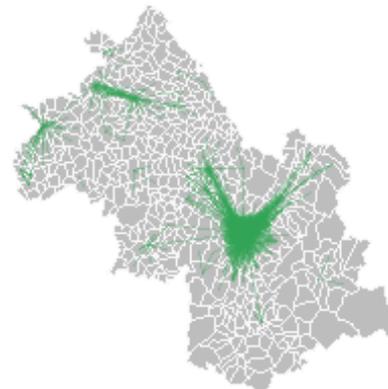
```
plot(fdc, col="#bdbdbd",border="#f7f7f7",lwd=0.5)
segments(tabflow_seuil$X1, tabflow_seuil$Y1, tabflow_seuil$X2,
```

Carte de Flux de Isere



(c) Prog R 2015

Carte de Flux de Isere



(c) Prog R 2015

```
tabflow_seuil$Y2,col="#2ca25f", lwd=tabflow_seuil$size)
title(main="Carte de Flux de Isere ", sub="(c) Prog R 2015", cex.sub=0.9)
```

2.3.4 Application d'une variable visuelle de Transparence aux figurés

L'application d'une transparence aux figurés passe par une action au niveau de la couleur (voir Chapitre 5), où l'on mobilise le canal alpha. Pour plus d'informations sur la définition de palettes de couleurs générales ou personnalisées, on peut consulter le billet suivant de N. Lambert (2015) : <http://neocarto.hypotheses.org/1458>.

```
dev.off() #nettoyer la fenêtre de graphique
## null device
##      1
add.alpha<-function(col,alpha=0.3)
{
  if (missing(col))
    stop ("choisissez un vecteur de couleur")
  apply(sapply(col,col2rgb)/255,2,function(x)
    rgb(x[1],x[2],x[3],alpha=alpha))
}
```

Tracé du fond de carte et des liens par rapport au seuil choisi

```
plot(fdc, col="#bdbdbd",border="#f7f7f7",lwd=0.5) # affichage du fond de carte

arrows(tabflow_seuil$X1, tabflow_seuil$Y1, tabflow_seuil$X2,
tabflow_seuil$Y2,col=add.alpha("#31a354"), lwd=tabflow_seuil$size,length=0.1,code=0)

title(main="Carte de Flux de Isere ", sub="(c) Prog R 2015", cex.sub=0.9)
```

2.3.5 Tracé de flèches

Pour tracer des flèches, il convient d'appliquer des « têtes » aux liens, en introduisant le paramètre code=2 (dans le sens Fij) ou code=1 dans le sens (Fji).

Les paramètres d'affichage sont les suivants :

```
# flèche de i vers j : code=2 et de j vers i, code=1
# double tete de fleche entre i et j : code=3
# pas de tete de fleche code=0
```

Cartographie en utilisant des flèches, présentant un seuil de transparence

```
#dev.off() #nettoyer la fenêtre de graphique

#plot(fdc, col="#bdbdbd",border="#f7f7f7",lwd=0.5)

#arrows(tabflow$X1, tabflow$Y1, tabflow$X2, tabflow$Y2,col=add.alpha("#31a354"),
lwd=tabflow$size,length=0.1,code=0)
```

3. Procédure de sélection des flux à cartographier, selon un critère spatial (de voisinage binaire, ordinal ou de distance parcourue)

Cette annexe décrit des procédures permettant de sélectionner les flux selon des critères liés à l'espacement des lieux, présentés dans le chapitre 7, à savoir :

- (1) le voisinage spatial binaire ($C_{ij}=1$) ;
- (2) le voisinage spatial ordinal (CK_{ij}) avec $k=1,2,\dots$, tels que présentés dans le Chapitre 7
- (3) le voisinage spatial des *k plus proches voisins* (présentée en complément des deux précédentes)
- (4) la distance parcourue (D_{ij}).

Le cas d'illustration est la matrice de l'Isère.

3.1 Préparation

3.1.1 Chargement des données numériques et spatiales (zones)

Ici la matrice (Fij) à étudier est celle de l'Isère

```
tab <-
read.csv2("D:/FRANCOISE/R/Isere/isere_Fij.csv",header=TRUE,sep=";",stringsAsFactors=FALSE,encoding="UTF-8",dec=".",check.names=FALSE)
Fij <- as.matrix(tab[,-1])
row.names(Fij) <- tab[,1]

class(Fij)
## [1] "matrix"
```

Chargement du fond de carte (fdc) composé de zones

```
fdc <- readShapeSpatial("./Isere_com.shp")

str(fdc)
class(fdc)
#pt<-readShapeSpatial("./Isere_coord.shp") # chargement d'un fichier de points de coordonnées (x,y)
```

3.1.2 Préparation d'un fichier de points (x,y)

Si le fichier de points de coordonnées n'est pas disponible, on le prépare comme ci-dessous.

Extraction des centroïdes des zones et création d'un fichier de Points (pt) comprenant leur code identifiant (CODE), leurs coordonnées en X et en Y.

```
pt <- cbind(fdc@data[, "INSEE_COM"], as.data.frame(coordinates(fdc)))
colnames(pt) <- c("CODE", "X", "Y")
head(pt)
##      CODE      X      Y
## 0 38226 917543.1 6416638
## 1 38206 933896.6 6464035
## 2 38301 907669.5 6413449
## 3 38448 840772.2 6484481
## 4 38287 893903.3 6483937
## 5 38153 904873.1 6458243
```

Réalisation d'une jointure attributaire entre la liste de flux (tabflow) et le fichier de points (pt)

```
tabflow = data.frame(tabflow, pt[match(tabflow[, "i"], pt[, "CODE"]), 2:3])
tabflow = data.frame(tabflow, pt[match(tabflow[, "j"], pt[, "CODE"]), 2:3])
```

```
colnames(tabflow) <- c("i", "j", "Fij", "X1", "Y1", "X2", "Y2")
head(tabflow)
##      i      j Fij      X1      Y1      X2      Y2
## 253 38001 38001  0 902347.2 6496924 902347.2 6496924
## 63  38002 38001  0 934589.3 6466585 902347.2 6496924
## 260 38003 38001  0 845671.2 6472517 902347.2 6496924
## 381 38004 38001  0 892928.9 6460737 902347.2 6496924
## 51  38005 38001  0 937063.9 6456949 902347.2 6496924
## 47  38006 38001  0 943664.5 6480414 902347.2 6496924
```

3.2 Construction d'une matrice de voisinage ordinaire (Cij=1)

Création de la matrice de contiguïté `contig` : (Cij=1) où la valeur 1 correspond à l'existence d'une frontière de longueur non nulle entre `i` et `j`, sinon (Cij=0).

```
# Création de la matrice de contiguïté d'ordre 1 (k=1)
# La fonction gintersect agit sur la géométrie du fdc

contig<-gIntersects(fdc, byid = TRUE, prepared=TRUE)
row.names(contig)<-fdc@data$INSEE_COM
colnames(contig)<-fdc@data$INSEE_COM

# mise à 1 ou à 0 de la matrice
for (i in 1:nrow(contig))
  for (j in 1:ncol(contig))
    {if (contig[i,j]==TRUE) {contig[i,j]<-1}
      if (contig[i,i]!=0) {contig[i,i]<-0}
    }
```

Transformation de `contig` en liste

```
#Transformation en Liste de contig
tab <-melt(contig)
names(tab) = c("CODE_i", "CODE_j", "cij")
head(tab)
##  CODE_i CODE_j cij
## 1  38226 38226  0
## 2  38206 38226  0
## 3  38301 38226  0
## 4  38448 38226  0
## 5  38287 38226  0
## 6  38153 38226  0
```

Nettoyage de la table en supprimant les valeurs nulles

```
# on renomme le tab en ordre_1 pour préparer l'export
ordre_1<-tab
# Suppression des valeurs égales à 0
ordre_1<-ordre_1[ordre_1[,"cij"]!=0,]
# export éventuel de la table
# write.table(ordre_1,file="CKij_1.csv",sep=";",eol="\n")
```

La représentation de cette matrice conduit à un graphe de voisinage d'ordre 1.

Pour cartographier les flux qui se produisent entre des zones adjacentes, il convient de réduire la matrice (Fij) en fonction de (Cij)=1. La procédure est la même que dans le cas d'une réduction des flux selon une métrique continue (voir section 3.5 ci après).

3.3 Construction d'une matrice de voisinage ordinal (CKij)

On construit ici une matrice de contiguïté d'ordre 2 : (CKij) avec $k=2$. Pour cela, on repart de la matrice précédente (Cij=1) car nous avons vu dans le chapitre 7 que (Cij=1) et (CKij) avec $k=1$ étaient identiques.

Nous construisons d'abord le voisinage d'ordre 2 puis le voisinage d'ordre 4

3.3.1 Construction de la matrice (CKij) avec $k=2$

L'objectif est de lancer une boucle grâce à laquelle on va rechercher, pour une entité (i) quelconque, le voisin d'ordre 1 de son voisin d'ordre 1 qui sera alors le voisin d'ordre 2 de (i).

```
# Création de La matrice d'ordre 2 à partir de La matrice d'ordre 1

CKij_1<-ordre_1
CKij_1$CODE_i1<-0
CKij_1$CODE_j1<-0
CKij_1$CODE_j2<-0
CKij_1$cij2<-0

for (k in 1:nrow(CKij_1))
{
  if (CKij_1$cij[k]==1)
  {
    col_i<-CKij_1$CODE_i[k]
    col_j<-CKij_1$CODE_j[k]

    for (v in 1:nrow(CKij_1))
    {
      if (CKij_1$CODE_i[v]==col_j && CKij_1$CODE_i[v]!=CKij_1$CODE_j[v])
      {
        CKij_1$CODE_i1[k]<-col_i #i ordre 1
        CKij_1$CODE_j1[k]<-col_j #j ordre 1
        CKij_1$CODE_j2[k]<- CKij_1$CODE_j[v] #j ordre 2
        CKij_1$cij2[k]<-2
      }
    }
  }
}
```

```
head(CKij_1)
##      CODE_i CODE_j cij CODE_i1 CODE_j1 CODE_j2 cij2
## 16  38366 38226  1  38366  38226  38456  2
## 96  38125 38226  1  38125  38226  38456  2
## 142 38321 38226  1  38321  38226  38456  2
## 154 38127 38226  1  38127  38226  38456  2
## 378 38403 38226  1  38403  38226  38456  2
## 444 38456 38226  1  38456  38226  38456  2
```

avec CODE_i = CODEi1 et CODEJ=CODE_J1 (à l'ordre 1) et CODE_J2= Lieu de destination à l'ordre 2. Les doublons permettent de vérifier l'opération.

On renomme les variables de la table précédente, avec (i) lieu d'origine et (j2) le lieu de destination à l'ordre 2, cij2 contient la modalité 2 choisie.

On supprime ensuite les valeurs nulles ou impossibles.

```
ordre_2<-data.frame(CKij_1$CODE_i,CKij_1$CODE_j2,CKij_1$cij2)
names(ordre_2) = c("i", "j2","cij2")

# Mise à zéro des Liaisons intra zones sur ckij ordre 2
for (z in 1:nrow(ordre_2))
```

```

if (ordre_2$i[z] == ordre_2$j2[z]) ordre_2$cij2[z]<-0
# Suppression des valeurs égales à 0
ordre_2<-ordre_2[ordre_2[, "cij2"]!=0,]
#write.table(ordre_2,file="./CKij_2.csv",sep=";",eol="\n")

```

La représentation de cette matrice conduit à un graphe de voisinage d'ordre 2. Utiliser pour cela le programme présenté en 2, pour obtenir la figure suivante.

Graphe de voisinage de l'Isère d'ordre 1



graphe de voisinage de l'Isère d'ordre 2



Remarque : pour cartographier les flux inférieurs à CKij avec k=2, il faut l'ajouter à l'ordre 1 pour conserver le principe de continuité spatiale.

3.3.2 Construction de la matrice (CKij) avec k=4

Nous reproduisons l'opération précédente en considérant que pour tout (i), le voisin d'ordre 2 de son voisin d'ordre 2 est le voisin d'ordre 4 de (i).

```

#Matrice d'ordre 4 à partir de La matrice d'ordre 2
ordre_2$CODE_i<-0
ordre_2$CODE_j2<-0
ordre_2$CODE_j4<-0
ordre_2$cij4<-0

for (k in 1:nrow(ordre_2))
{
  if (ordre_2$cij2[k]==2)
  {
    col_i4<-ordre_2$i[k]
    col_j4<-ordre_2$j2[k]

    for (v in 1:nrow(ordre_2))
    {
      if (ordre_2$i[v]==col_j4 && ordre_2$i[v]!=ordre_2$j2[v])
      {
        ordre_2$CODE_i[k]<-col_i4
        ordre_2$CODE_j2[k]<-col_j4
        ordre_2$CODE_j4[k]<-ordre_2$j2[v]
        ordre_2$cij4[k]<-4
      }
    }
  }
}

```

```

    }
  }
}

head(ordre_2)
##      i      j2 cij2 CODE_i CODE_j2 CODE_j4 cij4
## 1 38366 38456   2 38366  38456  38224   4
## 2 38125 38456   2 38125  38456  38224   4
## 3 38321 38456   2 38321  38456  38224   4
## 4 38127 38456   2 38127  38456  38224   4
## 5 38403 38456   2 38403  38456  38224   4
## 7 38005 38547   2 38005  38547  38397   4
ordre_4<-data.frame(ordre_2$CODE_i,ordre_2$CODE_j2,ordre_2$cij4)
names(ordre_4) = c("i", "j4","cij4")

# Mise à zéro des Liaisons intra zones sur ckij ordre 4
for (z in 1:nrow(ordre_4))
  if (ordre_4$i[z] == ordre_4$j4[z]) ordre_4$cij4[z]<-0

# Suppression des valeurs égales à 0
ordre_4<-ordre_4[ordre_4[, "cij4"]!=0,]

#write.table(ordre_4,file="./CKij_4.csv",sep=";",eol="\n")

```

La représentation de cette matrice conduit au graphe de voisinage suivant.

Graphe de voisinage de l'Isère, d'ordre 4



Remarque : pour cartographier cette table d'ordre 4 [(Fij) ← [(CKij, k<4)], il faut l'ajouter à l'ordre 1 et à l'ordre 2 pour conserver le principe de continuité spatiale. Il est également possible de ne visualiser que la sélection des flux qui franchissent 4 distances (et non de 1 à 4) [(Fij) ← [(CKij, k=4)].

3.4 Construction et représentation des *k plus proches voisins*

Cette section est inspirée de la vignette de R. Bivand (2015) sur la création des graphes de voisinage.

L'objectif pour nous est de construire une matrice de voisinage spatial qui n'est pas fondée sur le franchissement d'une limite de zone, comme dans les sections précédentes, mais sur la proximité des barycentres des zones.

Deux raisonnements sont possibles :

- (1) lorsque l'approche est ordinale et spatiale : elle considère la proximité des lieux en mobilisant leurs coordonnées (x,y);
- (2) lorsque est réticulaire : elle correspond à la distance géodésique, c'est-à-dire au nombre de liens qu'il convient de parcourir pour aller d'un point à un autre.

Pour plus d'informations, voir l'Annexe A.2.

3.4.1 Création du graphe de voisinage (*k voisins*)

Connexion au répertoire de travail et chargement des packages requis.

```
rm(list=ls())
setwd("C:/FRANCOISE/R/3_voisinage")

library(maptools)
library(sp)

## Checking rgeos availability: TRUE
library(spdep)
library(Matrix)
```

Chargement du fond de carte, et création d'un fichier de coordonnées (x,y) et d'un autre des ID

```
## [40] "trade"
fdc <- readShapeSpatial("C:/FRANCOISE/R/3_voisinage/isere_com.shp")

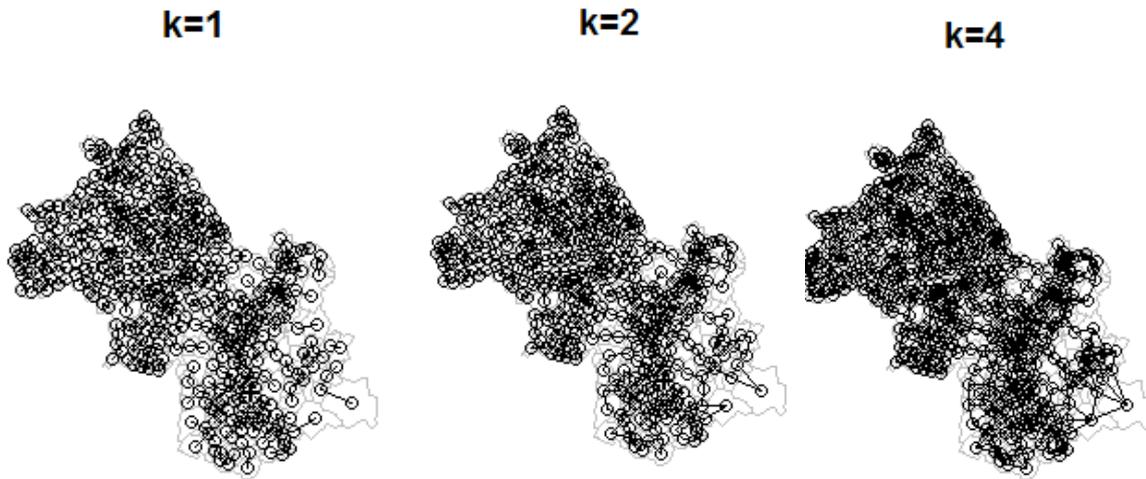
coords <- coordinates(fdc)
IDs <- row.names(as(fdc, "data.frame"))
```

Génération des graphes de voisinage à partir du fond de carte (fdc) d'ordre k=1 (fdc1_nb) à 4 : cette opération entraîne une visualisation immédiate du résultat

```
fdc1_nb <- knn2nb(knearneigh(coords, k = 1), row.names = IDs)
fdc2_nb <- knn2nb(knearneigh(coords, k = 2), row.names = IDs)
fdc4_nb <- knn2nb(knearneigh(coords, k = 4), row.names = IDs)
```

3.4.2 Visualisation des graphes de voisinages

Sortie cartographique brute



Cartographie mise en forme dans un logiciel de graphisme des graphes de voisinages (par simple suppression des cercles).



3.5 Construction d'une matrice de distances (Dij) et réduction de (Fij) en fonction de (Dij) restreinte

Cette section concerne la cartographie de flux qui parcourent une certaine classe de distance kilométrique. L'utilisation de (Dij) entraîne soit la réduction de (Fij) soit la réalisation d'une collection de cartes.

3.5.1 Chargement des données numériques

Préparation

```
rm(list=ls())

library(rgeos)
library(maptools)
library(proj4)
library(reshape2)
library(rCarto)
library(rgdal)

setwd("./distance")
```

Chargement de la matrice de flux (Fij) au format matriciel et transformation en liste.

```
Fij<-
read.csv2("./matrice_isere.csv",header=TRUE,sep=";",stringsAsFactors=FALSE,encoding="UTF-8",dec=".",check.names=FALSE)
str(Fij)
class(Fij)
## [1] "matrix"

matflow <- Fij
tabflow <-melt(matflow)
names(tabflow) = c("i", "j", "Fij")
```

Chargement du fond de carte et création d'une liste de points (pt) (x,y) à partir desquels nous allons créer une matrice de distances euclidiennes (Dij).

```
# Import et affichage du fond de carte
fdc <- readShapeSpatial("./isere_com.shp")
plot(fdc)
```



```
# Extraction des centroïdes
pt <- cbind(fdc@data[, "INSEE_COM"], as.data.frame(coordinates(fdc)))
colnames(pt) <- c("CODE", "X", "Y")
head(pt)
##      CODE      X      Y
## 0 38226 917543.1 6416638
## 1 38206 933896.6 6464035
## 2 38301 907669.5 6413449
## 3 38448 840772.2 6484481
## 4 38287 893903.3 6483937
## 5 38153 904873.1 6458243

# Jointure entre tabflow et pt
tabflow = data.frame(tabflow, pt[match(tabflow[, "i"], pt[, "CODE"]), 2:3])
tabflow = data.frame(tabflow, pt[match(tabflow[, "j"], pt[, "CODE"]), 2:3])
colnames(tabflow) <- c("i", "j", "Fij", "X1", "Y1", "X2", "Y2")
head(tabflow)
##      i      j Fij      X1      Y1      X2      Y2
## 253 38001 38001  0 902347.2 6496924 902347.2 6496924
## 63  38002 38001  0 934589.3 6466585 902347.2 6496924
## 260 38003 38001  0 845671.2 6472517 902347.2 6496924
## 381 38004 38001  0 892928.9 6460737 902347.2 6496924
## 51  38005 38001  0 937063.9 6456949 902347.2 6496924
## 47  38006 38001  0 943664.5 6480414 902347.2 6496924
```

3.5.2 Construction de la matrice de distance (Dij) et réduction de (Fij)

On définit au préalable le seuil de distance (dist) à appliquer à (Dij).

On prend l'exemple de 19 km comme variable de restriction de (Dij), pour ne représenter que les flux qui parcourent moins de 19 km.

```
# variable de restriction de La matrice Dij
#-----
dist=19000      #Indiquer une valeur
```

Construction de la matrice de distances, comme variable (Dij) de tabflow.

La variable (Dij_rest) correspond à la sélection des distances inférieures à 19km, sinon à 0

La variable (Dij_Fij) correspond à la sélection de (Fij) par rapport à (Dij_rest)

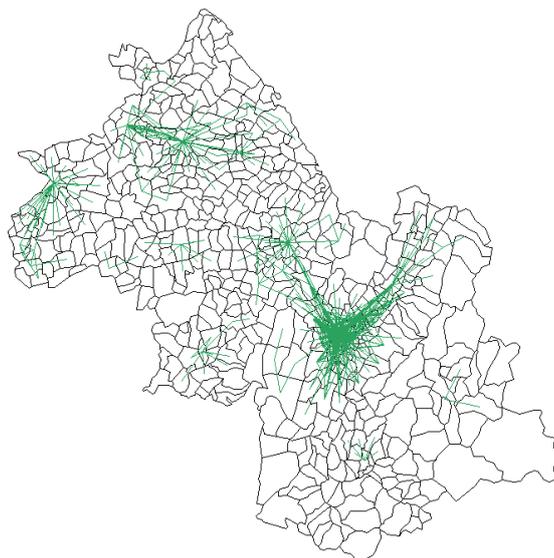
```
# Construction de La matrice de distances
for (k in 1:nrow(tabflow))
{
  tabflow$Dij[k] <- sqrt((tabflow$X1[k]-tabflow$X2[k])^2+(tabflow$Y1[k]-tabflow$Y2[k])^2)

  if (tabflow$Dij[k]<=dist) tabflow$Dij_rest[k] <- tabflow$Dij[k] # application de (dist)
  else tabflow$Dij_rest[k] <- 0
  if (is.na.data.frame(tabflow$Dij_rest[k])==TRUE) {tabflow$Dij_rest[k] <- 0}
  if (tabflow$Dij_rest[k]!=0) tabflow$Dij_Fij[k] <- tabflow$Fij[k] # restriction de (Fij) en
  fonction de (Dij)
  else tabflow$Dij_Fij[k] <- 0
  if (is.na.data.frame(tabflow$Dij_Fij[k])==TRUE) {tabflow$Dij_Fij[k] <- 0}
}
```

Cartographie de (Fij) \leftarrow [(Dij) > 19km]

```
maxsize <- 10 #epaisseur du plus gros Lien
tabflow$size_r <- (tabflow$Dij_Fij/max(tabflow$Dij_Fij))*maxsize
for (k in 1:nrow(tabflow))
{if (tabflow$size_r[k]>0.2)
  trace <- segments(tabflow$X1[k], tabflow$Y1[k], tabflow$X2[k],
tabflow$Y2[k], col="#2ca25f", lwd=tabflow$size_r[k])
}
```

Cartographie de flux locaux (parcourant moins de 19 km)



4. Procédure de sélection des flux à cartographier, selon un critère global

Cette annexe est très largement inspirée du programme *Flows Analysis Carto* de C. Grasland, (Version de Janvier 2011). Elle a été complétée en Janvier 2014 pour opérationnaliser les propositions du chapitre 8.

Ce programme permet ainsi d'enchaîner les traitements présentés dans le chapitre 8 à savoir :

- (1) le test de la (log) normalité d'une série de flux (Fij) ;
- (2) la représentation graphique de la distribution de fréquences simples et cumulées, sous la forme d'une courbe de Lorenz,
- (3) l'estimation aisée du double critère de sélection des flux, selon un *objectif de significativité* des valeurs ou de *densité* de la matrice (en nombre de liens représentés par rapport à l'information de flux totale).

Le cas d'application est la matrice de l'Isère.

4.1 Préparation

```
#purge de l'environnement de travail
rm(list=ls())
setwd("D:/FRANCOISE/R/Isere/")

# déclaration des packages requis
library(reshape2)
library(sp)
library(lattice)
```

4.1.1 Chargement de la matrice (Fij) sous forme de liste

Chargement de la matrice à étudier, ici sous la forme de liste (si elle ne l'est pas, utiliser la fonction melt). On notera que dans le tableau que l'on importe ci-après figure la distance kilométrique DIJ séparant les lieux d'origine et de destination et leurs coordonnées respectives. La variable TOT correspond à la partie entière de la variable FIJ.

```
tabflow <- read.table("isere_flux.csv", header=TRUE, sep=";", dec=",")
head(tabflow,5)
##  CODEi CODEj TOT      FIJ      DIJ Origine_X Origine_Y Dest_X Dest_Y
## 1 38001 38012 15 15.3317 60.16644 901700 6496300 903300 6502100
## 2 38001 38013 7 7.6658 168.75130 901700 6496300 894800 6480900
## 3 38001 38022 30 30.6634 111.00450 901700 6496300 899600 6507200
## 4 38001 38028 3 3.8329 31.01612 901700 6496300 902800 6493400
## 5 38001 38029 11 11.4988 65.05382 901700 6496300 897100 6500900
tab<-tabflow
head(tab)
##  CODEi CODEj TOT      FIJ      DIJ Origine_X Origine_Y Dest_X Dest_Y
## 1 38001 38012 15 15.3317 60.16644 901700 6496300 903300 6502100
## 2 38001 38013 7 7.6658 168.75130 901700 6496300 894800 6480900
## 3 38001 38022 30 30.6634 111.00450 901700 6496300 899600 6507200
## 4 38001 38028 3 3.8329 31.01612 901700 6496300 902800 6493400
## 5 38001 38029 11 11.4988 65.05382 901700 6496300 897100 6500900
## 6 38001 38053 45 45.9951 246.75899 901700 6496300 877600 6501600
```

4.1.2 Préparation des matrices complémentaires – forme liste

Les matrices complémentaires sont issues de la décomposition éventuelle de $(Fij)=f1$
 $Fji=f2$, $(Fij^+)=FSij$; $(Fij^-)=FDij$

```

i<-tab$CODEi
j<-tab$CODEj
Fij<-tab$TOT
head(tab)
##   CODEi CODEj TOT      FIJ      DIJ Origine_X Origine_Y Dest_X Dest_Y
## 1 38001 38012 15 15.3317 60.16644 901700 6496300 903300 6502100
## 2 38001 38013 7 7.6658 168.75130 901700 6496300 894800 6480900
## 3 38001 38022 30 30.6634 111.00450 901700 6496300 899600 6507200
## 4 38001 38028 3 3.8329 31.01612 901700 6496300 902800 6493400
## 5 38001 38029 11 11.4988 65.05382 901700 6496300 897100 6500900
## 6 38001 38053 45 45.9951 246.75899 901700 6496300 877600 6501600
f1<-data.frame(i,j,Fij)
f2<-data.frame(i,j,Fij)
names(f2)<-c("i","j","Fji")

f<-merge(f1,f2,by=c("i","j"),all.X=TRUE,all.Y=TRUE)
f$FSij<-f$Fij+f$Fji
#f$FDij<-f$Fji-f$Fij

head(f)
##      i      j Fij Fji FSij
## 1 38001 38012 15 15 30
## 2 38001 38013 7 7 14
## 3 38001 38022 30 30 60
## 4 38001 38028 3 3 6
## 5 38001 38029 11 11 22
## 6 38001 38053 45 45 90

```

4.1.3 Choix de la matrice à analyser (par exemple : Fij^+)

On choisit d'analyser le volume bilatéral de flux : $FSIJ = (Fij^+)$ que l'on recopie dans la matrice $Fij = (mytab)$. On crée ensuite une colonne supplémentaire indiquant l'existence d'un lien (link), la colonne flow indiquant la valeur du flux.

```

# on choisit FSIJ (volume bilatéral)
mytab<-data.frame(f$i,f$j,f$FSij)
names(mytab)<-c("i","j","flow")
mytab$link<-1
mytab<-mytab[mytab$flow>0,]

head(mytab,10)
##      i      j flow link
## 1 38001 38012 30 1
## 2 38001 38013 14 1
## 3 38001 38022 60 1
## 4 38001 38028 6 1
## 5 38001 38029 22 1
## 6 38001 38053 90 1
## 7 38001 38067 6 1
## 8 38001 38080 136 1
## 9 38001 38087 6 1
## 10 38001 38104 44 1
#tail(mytab,10)

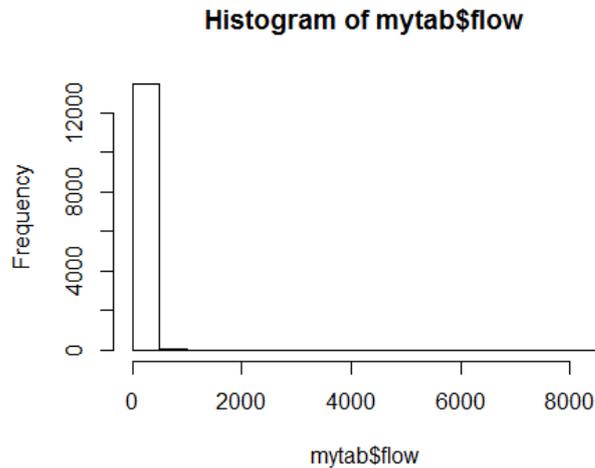
```

4.2 Analyse statistique de la normalité de (Fij)

4.2.1 Histogramme de fréquences

On trace l'histogramme de (Fij) afin de vérifier son allure par rapport à la Loi normale

```
#par(mfrow=c(1,1))
hist(mytab$flow)
```



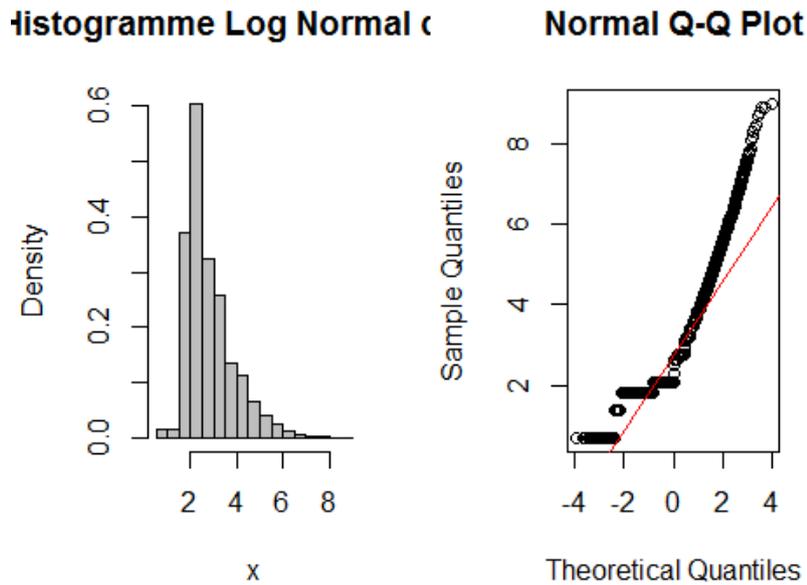
L'histogramme précédent étant fortement dissymétrique à gauche, on teste ensuite sa log-normalité. Pour cela, on extrait le résumé statistique de la série, on trace l'histogramme de distribution de $\log(\text{Fij})$ et le diagramme Quantile-Quantile.

4.2.2 Diagramme Quantile-Quantile

```
x<-log(mytab$flow)

summary(x)
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.6931 2.0790 2.3030 2.7650 3.3320 8.9880
mean(x)
## [1] 2.764878
sd(x)
## [1] 1.080829
par(mfrow=c(1,2))
hist(x,prob=TRUE,col="grey",
     main="Histogramme Log Normal de Fij")

qqnorm(x);qqline(x, col = 2)
```



L'ajustement log-normal étant imparfait, on décide de réaliser une analyse de concentration des valeurs de flux en fonction des liens, en mobilisant l'indice de Gini et la courbe de Lorenz.

4.3 Analyse de la concentration des valeurs de flux

4.3.1 Calcul de l'indice de Gini

On crée au préalable deux nouvelles variables dans la matrice $F_{ij}=(mytab)$: `flowcum` pour les valeurs cumulées des flux et `linkcum` pour celle des liens, qui sont triées par ordre décroissant des valeurs cumulées des flux ET celle des liens.

```
mytab$flowcum<-cumsum(mytab$flow)/sum(mytab$flow)
mytab$linkcum<-cumsum(mytab$link)/sum(mytab$link)
head(mytab)
##          i      j flow link  flowcum  linkcum
## 4682  38185 38421 8004    1 0.01385125 7.366483e-05
## 3619  38151 38185 7358    1 0.02658457 1.473297e-04
## 10000 38421 38185 7352    1 0.03930751 2.209945e-04
## 4592  38185 38151 6314    1 0.05023414 2.946593e-04
## 4093  38169 38185 5828    1 0.06031973 3.683241e-04
## 4617  38185 38229 4814    1 0.06865056 4.419890e-04
```

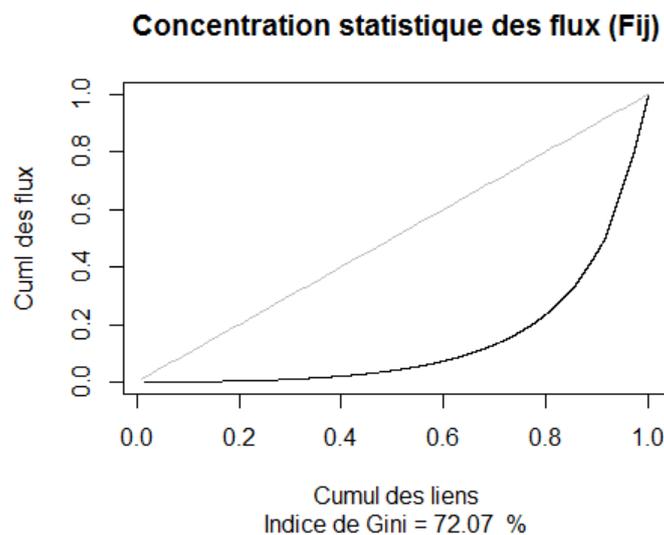
On calcule ensuite l'indice de Gini correspondant à la part d'information de flux transmise par la part de liens représentés.

```
gini<-function(vec1,vec2)
{
  tot<-vec2[1]/2*vec1[1]
  i<-2
  while(i <=length(vec1))
  {tot<-tot+(vec1[i]-vec1[i-1])*(vec2[i]+vec2[i-1])/2
  i<-i+1
  }
  res<-2*(0.5-tot)
  res
}
indice<-gini(mytab$flowcum,mytab$linkcum)
```

4.3.2 Courbe de concentration des valeurs de flux

On trace enfin la courbe de concentration statistique des flux correspondante (Fij/Lij)

```
par(mfrow=c(1,1))
title<-paste("Indice de Gini =",round(indice*100,2)," %")
plot(mytab$linkcum~mytab$flowcum, type="l",
     main="Concentration statistique des flux (Fij)",
     sub=title,
     xlab="Cumul des liens",
     ylab="Cuml des flux")
lines(1:100/100,1:100/100,col="grey")
```



Nous avons ensuite deux possibilités de sélection de l'information de flux à cartographier : selon significativité de l'information (Fij) *versus* selon la densité de liens (Lij) représentée.

4.4 Sélection globale selon un objectif de *significativité* de l'information

On souhaite représenter 80% de l'information totale de flux.

On définit pour cela un seuil de 80% (critflow=0.80).

```
critflow<-0.80
sel<-mytab[mytab$flowcum<critflow,]
x<-tail(sel,1)
seuil<-x[3]
critlink<-x[6]
resul<- paste("seuil =", round(seuil,0)," ---",
             "flows =", round(critflow*100,2),"% ---",
             "links=",round(critlink*100,2),"%")
resul
## [1] "seuil = 30 --- flows = 80 % --- links= 23.45 %"
```

Le résultat obtenu est le suivant :

- seuil= valeur de flux ;
- flows= part en % de l'information totale ;
- links = part de liens en % du total de liens

Ainsi, la cartographie des valeurs de flux supérieures à 30 correspond à 80% de l'information totale (critère de significativité) et à 23,4% des liens (critère de densité).

4.5 Sélection globale selon un objectif de *densité* de la matrice

On souhaite représenter 0,2% du total des liens. On définit pour cela le nombre de liens à représenter (critlink=0.002).

```
critlink<-0.002
sel<-mytab[mytab$linkcum<critlink,]
x<-tail(sel,1)
x
##          i      j flow link  flowcum  linkcum
## 4943 38193 38449 1604    1 0.1599435 0.00198895
seuil<-x[3]
critflow<-x[5]
resul<- paste("seuil =", round(seuil,0)," ---",
              "flows =", round(critflow*100,2),"% ---",
              "links=",round(critlink*100,2),"%")
resul
## [1] "seuil = 1604 --- flows = 15.99 % --- links= 0.2 %"
```

Le résultat obtenu est le suivant :

- seuil=valeur de flux ;
- flows= part en % de l'information totale de flux ;
- links = part de liens en % du total de liens

Ainsi, la cartographie de 0,22% de la densité de la matrice correspond à la représentation des valeurs de flux supérieures à 1604, qui contribuent à 15,99% de l'information totale.

5. Procédure de sélection des flux à cartographier, selon un critère local (« flux dominants »)

Cette annexe permet de sélectionner les flux selon des critères locaux, présentés dans le Chapitre 9, en amont de leur cartographie:

- (1) des flux majeurs $\text{Max}_i(\text{Fij})$: la procédure est alors identique à la précédente, car l'on sélectionne les flux qui répondent à un critère unique ; la différence réside dans la focale, c'est-à-dire dans la partie de la matrice mobilisée pour l'application de ce critère (ici les lieux d'origine) correspondant alors à $(\text{Fij}) \leftarrow \text{Max}_i(\text{Fij})$
- (2) et dominants [... et $(\text{Dj}) > (\text{Di})$].

Le cas d'illustration est celui des flux commerciaux internationaux de la base de données Chelem ou des navettes de l'Isère.

5.1 Préparation

5.1.1 Connexion au répertoire de travail et chargement des packages requis.

```
#purge environnement de travail
rm(list=ls())
setwd("D:/FRANCOISE/R/chelem/")

# declaration des packages requis
library(reshape2)
library(gdata)
library(rCarto)
library(maptools)
library(RColorBrewer)
library(classInt)
```

5.1.2 Chargement de la matrice (Fij) à étudier

On charge ici les flux commerciaux internationaux.

```
tab <-
read.csv2("./CHELEM_Fij_mat.csv",header=TRUE,sep=";",stringsAsFactors=FALSE,encoding="UTF-8",dec=",")
```

5.2 Sélection de flux majeurs et dominants (méthode classique Nuysten-Dacey, 1961)

Chargement de la matrice et vérification du format de l'objet

```
mij <- as.matrix(tab[,-1])
row.names(mij) <- tab[,1]
class(mij)
## [1] "matrix"
```

Création d'une Fonction de mise à zéro de la diagonale principale

```
F_zero<-function(x)
{
  nbi<-dim(mij)[1]
  nbj<-dim(mij)[2]

  for (i in 1:nbi)
    for (j in 1:nbj)
```

```

    if (i==j)x[i,j]<-0
  resul<-x
}
Fij<-F_zero(mij)

```

Calcul des matrices complémentaires issues de la décomposition de (Fij) :

Fji, Fij_plus=(Fij⁺) ; Fij_moins=(Fij⁻)

```

Fji<-t(Fij)
Fij_plus<-Fij+Fji
Fij_moins<-Fij-Fji

```

Création d'une Fonction temporaire, qui permet de garder la structure de (Fij) (les codes identifiants) ; elle sera utilisée tout au long du programme.

```

Tampon<-function(x)
{
  nbi<-dim(Fij)[1]
  nbj<-dim(Fij)[2]
  Temp_1<-Fij
  for (i in 1:nbi)
    for (j in 1:nbj)
      Temp_1[i,j]<-x
  Temp_1
}
#Mise à zéro de la fonction Tampon
Temp<-as.matrix (Tampon (0))

```

5.2.1 Création et application d'une fonction de sélection des Flux majeurs

Création d'une fonction : (F_MAJ)

```

#Fonction flux majeur
#-----
F_MAJ<-function(Zij)
{
  nbi<-dim(Fij)[1]
  nbj<-dim(Fij)[2]

  MAX<-apply(Zij,1,max)

  for (i in 1:nbi)
    for (j in 1:nbj)
    {
      if (Zij[i,j]==MAX[i]) Temp[i,j]<-Zij[i,j]
    }
  resul<-Temp
}

```

Application de la F_MAJ aux différentes matrices

Export éventuel de la matrice réduite au format csv pour utilisation ultérieure. Choix de la matrice avec laquelle poursuivre (Fij), dans notre cas

```

Flux_MAJ_Fij<-F_MAJ(Fij)
Flux_MAJ_Fji<-F_MAJ(Fji)
Flux_MAJ_Fij_plus<-F_MAJ(Fij_plus)
Flux_MAJ_Fij_moins<-F_MAJ(Fij_moins)

#write.table(Flux_MAJ_Fij,file="./Flux_MAJ_Fij.csv",sep=";",eol="\n")
#write.table(Flux_MAJ_Fji,file="./Flux_MAJ_Fji.csv",sep=";",eol="\n")
#write.table(Flux_MAJ_Fij_plus,file="./Flux_MAJ_Fij_plus.csv",sep=";",eol="\n")
#write.table(Flux_MAJ_Fij_moins,file="./Flux_MAJ_Fij_moins.csv",sep=";",eol="\n")

```

```
#suppression variable temporaire
rm(Temp,mij)
```

5.2.2 Création et application d'une Fonction flux dominant

Création de (F_DOM)

```
Temp<-as.matrix(Tampon(0))
F_DOM<- fonction(Pij)
{
  nbi<-dim(Fij)[1]
  nbj<-dim(Fij)[2]

  SUM<-apply(Pij,2,sum)

  #Fonction flux majeur
  MAX<-apply(Pij,1,max)

  for (i in 1:nbi)
    for (j in 1:nbj)
      {
        if (Pij[i,j]==MAX[i] && SUM[i]<SUM[j]) Temp[i,j]<-Pij[i,j]
      }
  resul<-Temp
}
```

Choix de la matrice à réduire aux flux dominants, à laquelle appliquer la F_DOM – export éventuel en fichier CSV, pour utilisation ultérieure et choix de la matrice qui sera cartographiée : (Fij) dans notre cas

```
Flux_DOM_Fij<-F_DOM(Fij)
Flux_DOM_Fji<-F_DOM(Fji)
Flux_DOM_Fij_plus<-F_DOM(Fij_plus)
Flux_DOM_Fij_moins<-F_DOM(Fij_moins)

#write.table(Flux_DOM_Fij,file="./Flux_DOM_Fij.csv",sep=";",eol="\n")
#write.table(Flux_DOM_Fji,file="./Flux_DOM_Fji.csv",sep=";",eol="\n")
#write.table(Flux_DOM_Fij_plus,file="./Flux_DOM_Fij_plus.csv",sep=";",eol="\n")
#write.table(Flux_DOM_Fij_moins,file="./Flux_DOM_Fij_moins.csv",sep=";",eol="\n")

#suppression variable temporaire
rm(Temp,mij)
## Warning in rm(Temp, mij): objet 'mij' introuvable
```

5.3 Sélection de flux dominants – variante sur le critère (1) : (i) envoi x% à (j) et (Dj)>(Di)

5.3.1 Création et application d'une fonction de calcul de fréquences

Création d'une Fonction Frequence (Freq) de calcul de fréquence en ligne (Fij/Oi) et création d'une variable temporaire

```
Temp<-as.matrix(Tampon(0))

# Création d'une fonction Freq de calcul de fréquence en ligne (Fij/Oi)et création d'une
variable temporaire
#-----
Freq<-function(Sij)
{
  nbi<-dim(Sij)[1]
```

```

nbj<-dim(Sij)[2]

SUM<-apply(Sij,1,sum)

for (i in 1:nbi)
  for (j in 1:nbj)
  {
    Temp[i,j]<-(Sij[i,j]/SUM[i])*100
  }
resul<-Temp
}

```

Application de la fonction Freq sur (Fij) et (Fji) et création de nouvelles matrices

```

Freq_1<-Freq(Fij)
Freq_1_Fji<-Freq(Fji)

```

Suppression de la variable temporaire

```
rm(Temp)
```

5.3.2 Application de la fonction flux dominant sur variante critère 1 exprimée en ($\alpha\%$)

Attention : Une variante classique consiste à interpréter les *flux majeurs* comme ceux qui correspondent à plus de ($\alpha\%$) du total des flux émis par i : $\text{Max}_i(\text{Fij})$ exprimée sous une forme décimale ou en pourcentage %, correspondant par exemple à $(\text{Freq}_1)=(\text{Fij}/\text{O}_i)$. Cette variante est souvent exprimée ainsi : « (i) envoie 25% de ses flux vers (j) »

```

#Temp<-as.matrix(Tampon(0))

#Freq<-function(Sij)
#{
#  nbi<-dim(Sij)[1]
#  nbj<-dim(Sij)[2]

#  SUM1<-apply(Sij,1,sum)

#  for (i in 1:nbi)
#    for (j in 1:nbj)
#    {
#      Temp[i,j]<-(Sij[i,j]/SUM1[i])*100
#    }
#  resul<-Temp
#}
#Freq_1<-Freq(Fij)
#suppression variable temporaire
#rm(Temp)

```

Cette solution n'est pas développée car elle est impossible à cartographier en l'état. La valeur générée étant une donnée de rapport et non un stock, il n'est pas possible de l'utiliser pour paramétrer la largeur graphique du lien symbolisant le flux.

Pour utiliser cette matrice (Freq_1) il faut plutôt s'en servir pour réduire (Fij) , comme dans la section 5.4.

5.3.3 Application de la fonction flux dominant sur variante critère 2

Application du second critère sur (Fij) « ... et $D(i)$ inférieur à $D(j)$ » sur Fij et non sur la matrice (Freq_1) précédente.

Attention : cette matrice de fréquence comporte, en effet, des données sous forme décimale (ou en pourcentage), dont la cartographie directe est impossible. C'est pour cela qu'on réduira plutôt (Fij)=Fluxtemp en fonction de la sélection correspondant à (Freq_1), ce qui nous permettra ensuite de cartographier normalement et de manière adéquate les valeurs de flux.

```
Temp<-as.matrix(Tampon(0))

nbi<-dim(Fij)[1]
nbj<-dim(Fij)[2]
SUM<-apply(Fij,2,sum)

for (i in 1:nbi)
  for (j in 1:nbj)
  {
    if (SUM[i]<SUM[j]) Temp[i,j]<-Fij[i,j]
  }

Flux_temp<-Temp
rm(Temp)
```

5.3.4 Variable de restriction de (Fluxtemp) en fonction de (fluxmajvar1)

```
Temp<-as.matrix(Tampon(0))

nbi<-dim(Fij)[1]
nbj<-dim(Fij)[2]

for (i in 1:nbi)
  for (j in 1:nbj)
  {
    if (flux_dom_crit1_Fij[i,j]!=0) Temp[i,j]<-Flux_temp[i,j]
  }
Flu_DOM_var_crit1<-Temp

#write.table(Flu_DOM_var_crit1,file="FLU_DOM_crit1.csv",sep=";",eol="\n")
#suppression de la variable temporaire
rm(Temp)
```

5.4 Sélection de flux dominants – variante sur le critère (2) : « ... et le volume de flux reçu par (j) est supérieur à celui reçu par (i) »

Le volume de flux aux lieux correspondant à la somme des marges (O_i+D_j) de chacun des lieux. Le programme ci-dessous applique la variation des deux critères de sélection, de la manière suivante :

- (1) matrice de sélection de flux majeurs
- (2) « $(O_i+D_i)<(O_j+D_j)$ » avec $(O_i + D_j)$ correspondant au volume total de flux aux lieux.

Seule l'application du second critère est présentée.

5.4.1 Variation du second critère sur (Fij) « ... et (O_i+D_i) inférieur (O_j+D_j) »

```
Temp<-as.matrix(Tampon(0))

nbi<-dim(Fij)[1]
nbj<-dim(Fij)[2]

SUM1<-apply(Fij,1,sum)
SUM2<-apply(Fij,2,sum)
vo1<-SUM1+SUM2
```

```

for (i in 1:nbi)
  for (j in 1:nbj)
  {
    if (vol[i]<vol[j]) Temp[i,j]<-Fij[i,j]
  }

Flux_temp_vol<-Temp
rm(Temp)

```

5.4.2 Variable de restriction de fluxmajvar1 en fonction de flux_temp

```

Temp<-as.matrix(Tampon(0))

nbi<-dim(Fij)[1]
nbj<-dim(Fij)[2]

for (i in 1:nbi)
  for (j in 1:nbj)
  {
    if (Flux_MAJ_var1[i,j]!=0) Temp[i,j]<-Flux_temp_vol[i,j]
  }
Flux_DOM_var_crit1_crit2<-Temp

#write.table(Flux_DOM_var_crit1_crit2,file="Flux_DOM_var_crit1_crit2.csv",sep=";",eol="\n")
#suppression variable temporaire
rm(Temp)

```

Il est ensuite possible de cartographier les différentes matrices obtenues.

5.5 Sous Programme de cartographie d'une des matrices de flux dominants

Ce programme correspond à peu de choses près à ceux qui sont présentés dans les autres sections, il est adapté à l'utilisation d'une des matrices présentées dans l'espace RStudio Environnement global et ne comporte pas de sélection des valeurs de flux : la matrice de flux majeurs ou dominants qu'il faut cartographier correspondant déjà à une forme de sélection suffisante.

Les matrices créées précédemment ayant un format matriciel, il est nécessaire de les transformer en liste pour pouvoir les cartographier.

```

matflow <- flux
tabflow <-melt(matflow)
names(tabflow) = c("CODEi", "CODEj", "Fij")
head(tabflow)
##  CODEi CODEj      Fij
## 1     A     A      0.0
## 2     B     A 287742.2
## 3     C     A      0.0
## 4     D     A      0.0
## 5     E     A      0.0
## 6     F     A      0.0

```

Import et affichage du fond de carte au format shapefile (*.shp)

```

fdc <- readShapeSpatial("D:/FRANCOISE/R/chelem/CHELEM_in96.shp")

```

Extraction des centroïdes des zones et création d'un fichier de Points (pt) comprenant leur code identifiant, leurs coordonnées en X et en Y.

```

pt <- cbind(fdc@data[, "CHELEM_COD"], as.data.frame(coordinates(fdc)))
colnames(pt) <- c("CODE", "X", "Y")
head(pt)
##  CODE      X      Y
## 0  QEZ  7618035  881712.8

```

```
## 1 SG 1873574 -5095555.7
## 2 OBY 5029680 -3899413.9
## 3 NDA -12685642 -5781969.0
## 4 RAF 3904001 -3812535.7
## 5 MAA 9013024 8868008.0
```

Réalisation d'une jointure entre la liste de flux (tabflow) et le fichier de points (pt)

```
tabflow= data.frame(tabflow, pt[match(tabflow[, "CODEi"], pt[, "CODE"]),2:3])
tabflow = data.frame(tabflow, pt[match(tabflow[, "CODEj"], pt[, "CODE"]),2:3])
colnames(tabflow) <- c("i", "j", "TOT", "X1", "Y1", "X2", "Y2")
head(tabflow)
##   i j      TOT      X1      Y1      X2      Y2
## 70 A A  0.0 -5477939.2  871384.7 -5477939 871384.7
## 90 B A 287742.2 -3435874.5  467919.6 -5477939 871384.7
## 32 C A  0.0  207775.2 -4826863.9 -5477939 871384.7
##  7 D A  0.0  371352.9 -4375850.9 -5477939 871384.7
## 24 E A  0.0  782512.2 -4282521.0 -5477939 871384.7
## 43 F A  0.0 1054822.2 -5037381.9 -5477939 871384.7
```

Variables permettant de contrôler l'affichage des figurés : on dimensionne la *largeur graphique* maximale des bandes et on indique que seuls les liens existants (et non théoriques) seront représentés ; dans le cas d'une cartographie de (Fij)=(Fij_moins), cela signifie que seules les valeurs positives seront représentées.

```
nb<-0
tabflow<-tabflow[tabflow$TOT>nb,]
```

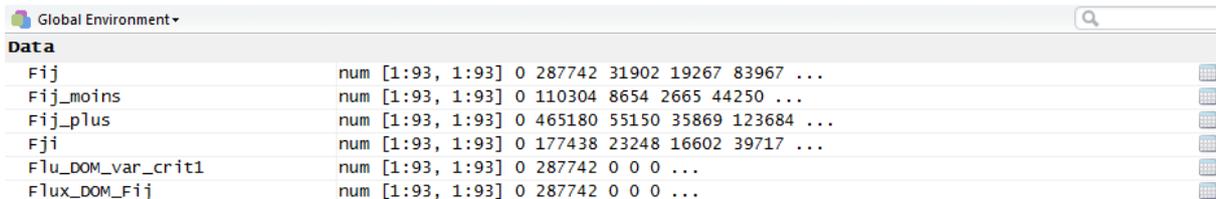
Paramètre de Taille graphique des flux (*largeur* du figuré) :

```
maxsize<-15 # épaisseur graphique du plus gros flux
tabflow$size<-(tabflow$TOT/max(tabflow$TOT))*maxsize
```

A noter ici que les flux ne font pas l'objet d'une sélection puisqu'ils le sont déjà (seuls les flux dominants peuvent être représentés ici). Si cela n'est pas le cas, il faudrait les sélectionner et pour cela, on choisirait alors, à titre d'exemple, de ne représenter que les flux supérieurs à la valeur moyenne... On commencerait par extraire le résumé statistique, ...

6. Procédure de cartographie directe d'une variable de l'environnement global de Rstudio©

Plutôt que de reproduire à chaque fois toute la procédure, il est possible d'optimiser davantage la production de cartes de flux en mobilisant la fenêtre *Environnement global* de l'interface RStudio© illustré ci-après qu'on appellera ensuite en source.



Global Environment	
Data	
Fij	num [1:93, 1:93] 0 287742 31902 19267 83967 ...
Fij_moins	num [1:93, 1:93] 0 110304 8654 2665 44250 ...
Fij_plus	num [1:93, 1:93] 0 465180 55150 35869 123684 ...
Fji	num [1:93, 1:93] 0 177438 23248 16602 39717 ...
Flu_DOM_var_crit1	num [1:93, 1:93] 0 287742 0 0 0 ...
Flux_DOM_Fij	num [1:93, 1:93] 0 287742 0 0 0 ...

L'objectif est de calculer dans un premier temps toutes les matrices dont on peut avoir besoin puis de sélectionner celle que l'on veut cartographier.

Comme précédemment, nous faisons appel à deux programmes distincts : le premier permet de préparer / de sélectionner les données et le second vise à les cartographier la matrice correspondante sur une carte.

Le cas d'application correspond aux flux commerciaux de marchandises (OMC, 2007) et la méthode de sélection est celle des flux majeurs et dominants (version classique Nyusten-Dacey (1961) selon une *logique de flux*)

6.1 Préparation

```
#purge environnement
rm(list=ls())
setwd("D:/FRANCOISE/R/OMC/")

# declaration des packages requis
library(reshape2)
library(gdata)
library(rCarto)
library(maptools)
library(RColorBrewer)
library(classInt)
```

Chargement de la matrice à étudier

```
tab <-
read.csv2("D:/FRANCOISE/R/OMC/OMC_Fij_mat.csv",header=TRUE,sep=";",stringsAsFactors=FALSE,encoding="UTF-8",dec=",")
# La matrice est ici au format matriciel ; si elle ne l'est pas, il faut la transformer en utilisant la fonction melt (présentée au début du programme de Cartographie - Voir Annexe R-2).
```

6.2 Chargement de l'environnement global

```
# Ces variables sont issues du programme de sélection selon des critères locaux :
source("D:/FRANCOISE/R/OMC/critere_locaux.R")
# Le programme critere_locaux permet de calculer Les flux majeurs & dominants selon la méthode Nyusten-Dacey ainsi que quelques variantes de la condition (1) et/ou (2)
```

Ces variables (*Data*) apparaissent dans la fenêtre Global Environnement de RStudio, comme illustré ci-dessous. Elles sont automatiquement chargées à la suite de la commande précédente qui lance le démarrage du code du programme Critères Locaux.

...

Ces variables générées automatiquement correspondent aux différentes matrices de flux disponibles pour la cartographie, leur descriptif est présenté dans le paragraphe suivant.

6.2.1 Liste des variables de l'environnement

```
#Fij : matrice de flux de (i) vers (j), par défaut asymétrique
#Fji : matrice transposée de Fij
#Fij_moins : matrice (Fij-) de la partie antisymétrique de (Fij)
#Fij_plus : matrice (Fij+) de la partie symétrique de (Fij)
#Flu_DOM_var_crit1 : variation critère 1 méthode NUYSTEN-DACEY (i envoie plus x% vers j ;
critère 2 inchangé) - matrice temporaire
#Flux_DOM_Fij : matrice de Flux dominant méthode NUYSTEN-DACEY sur (Fij)
#Flux_DOM_FIj_moins : matrice de Flux dominant méthode NUYSTEN-DACEY sur (Fij-)
#Flux_DOM_FIj_plus : matrice de Flux dominant méthode NUYSTEN-DACEY sur (Fij+)
#Flux_DOM_Fji : matrice de Flux dominant méthode NUYSTEN-DACEY sur Fji
#Flu_DOM_var_crit1_crit2 : matrice de flux dominant avec variation critères 1 & critère 2
méthode NUYSTEN-DACEY (i envoie plus x% vers j et le volume (Oi+Dj) de flux reçu par j est
supérieur à celui de i)
#Flux_MAJ_Fij : matrice de Flux majeur méthode NUYSTEN-DACEY sur Fij (=max(Fij))
#Flux_MAJ_FIj_moins : matrice de Flux majeur méthode de NUYSTEN-DACEY sur (Fij-)
#Flux_MAJ_FIj_plus : matrice de Flux majeur méthode NUYSTEN-DACEY sur (Fij+)
#Flux_MAJ_Fji : matrice de Flux majeur méthode NUYSTEN-DACEY sur (Fji)
#Flux_MAJ_var1 : matrice temporaire identique à
#Flux_temp : restriction de (Fij) au second critère méthode NUYSTEN-DACEY sur Fij (la
somme des flux reçus par j est supérieure à celle de i)
#Flux_temp_vol : restriction de (Fij) au second critère méthode NUYSTEN-DACEY sur Fij
(volume des flux reçus par j est supérieur à i)
#Freq_1 : matrice de transformation de (Fij) en pourcentage en ligne
#Freq_1_Fji : matrice de transformation de (Fji) en pourcentage en ligne
#Flux : variable(matrice) à cartographier appelée dans le programme Cartof
#flux_dom_crit1_Fij : matrice de Flux dominant avec variation critère 1 méthode NUYSTEN-
DACEY (i envoie plus x% vers j)
#flux_dom_crit1_Fji : Flux dominant avec variation critère 1 méthode NUYSTEN-DACEY (j
envoie plus x% vers i)
```

Le programme précédent calcule les flux dominants et leurs variantes pour les matrices (Fij), (Fji), (Fij⁺) et (Fij⁻). Le détail du script est disponible dans l'Annexe – R. 5. Suggestion de programmes de sélection de flux à cartographier selon un critère local

6.2.2 Cartographie d'une matrice située dans le "Global Environnement"

La procédure consiste à choisir une des matrices pré-calculée puis à lui appliquer la variable Flux qui entraîne sa cartographie.

```
# Dans cet exemple, la variable Flux entraîne l'exécution du programme de cartographie #
réalisée pour les données Chelem - Carto_flux_chelem
flux<-Flux_DOM_Fij

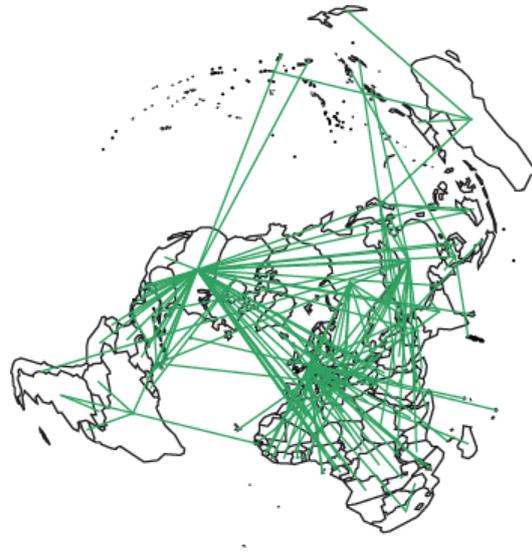
source("D:/FRANCOISE/R/OMC/Carto_chelem.R")
```

L'objectif est de réaliser sous R une carte telle que celles-ci qui décrivent les *flux dominants* ou les liens issus des *flux dominants*.

Carte de flux dominants de l'Isère



Carte de liens issue de flux dominants de l'Isère



Dès lors que la carte a été tracée sous Rstudio, il convient ensuite de l'exporter dans un logiciel de graphisme afin de finaliser sa mise en page.

7. Références

Les références suivantes ont été mobilisées pour les programmes de ces annexes, ou mentionnées dans le texte.

Bahoken, F., Chapron, P., *Cartographie de flux*, 5^e session du groupe ElementR [En ligne] <http://elementr.hypotheses.org/204> [lien vérifié le 06/09/2015].

Bivand, R., 2015, *Creating neighbours*, Vignette from Bivand, R. S., Pebesma, E., Gómez-Rubio, 2008, *Applied Spatial Data Analysis with R*, first edition, Springer-Verlag, New York. pp. 239–251. [En ligne] <https://cran.r-project.org/web/packages/spdep/vignettes/nb.pdf> [lien vérifié le 10/09/2015].

Casajus, N., *Analyses spatiales sous R*, Université du Québec, Département de biologie, chimie et géographie, 99 p.

Commenges, H., 2015, *RStudio, sa vie, son œuvre, ses ressources*. [En ligne] <http://elementr.hypotheses.org/237> [lien vérifié le 02/08/2015].

Giraud, T., Lambert, N., 2015, *Commented Scripts to Build Maps with Rcartography*, Vignette du package *Cartography*. [lien vérifié le 13/10/2015] <https://cran.r-project.org/web/packages/cartography/vignettes/cartography.html>

Grasland C., 2014, *Flows Analysis Carto* (Version de Janvier 2011, complétée en Janvier 2014), non publié.

Grasland C., Beauguitte, L., 2015, *Analyse de flux et modèles gravitaires*, 3^e session du groupe ElementR. [En ligne] <http://elementr.hypotheses.org/143> [lien vérifié le 06/09/2015].

Groupe ElementR, 2014, *R et Espace. Traitement de l'information géographique*, Framabook, 244 p. [En ligne] <http://framabook.org/r-et-espace/> [lien vérifié le 02/08/2015].

Lambert, N. (2015) [R], *Faire une carte de discontinuités avec R*, Billet de blog Néocarto [En ligne] <http://neocarto.hypotheses.org/1458> [lien vérifié le 02/08/2015].

Lambert, N. (2015) [R], *Créer des palettes de couleurs avec R*, Billet de blog Néocarto [En ligne] <http://neocarto.hypotheses.org/423> [lien vérifié le 02/08/2015].

8. Plan détaillé

1. Procédure de transformation d'une matrice de flux fermée non carrée en une matrice carrée	2
1.1 Justification	2
1.2 Préparation.....	3
1.2.1 Connexion au répertoire de travail et chargement des packages requis.	3
1.2.2 Chargement de la liste des identifiants	3
1.3 Construction d'une matrice vide et transformation de la matrice non carrée (format matriciel)	3
1.3.1 Construction d'une matrice vide	3
1.3.2 Transformation de cette matrice au format liste	3
1.3.3 Chargement de la matrice des données de flux observées non carrée (au format liste) ..	4
1.3.4 Transformation des données de flux observées en matrice carrée	4
2. Procédure de cartographie d'une matrice de flux (approche matricielle).....	6
2.1 Préparation.....	6
2.1.1 Chargement des données numériques.....	6
2.1.2 Préparation des matrices décomposées (F_{ij}^+), (F_{ij}^-) – forme matricielle.....	7
2.1.3 Chargement des données cartographiques (zones et/ou points)	7
2.2 Cartographie des liens (L_{ij}).....	8
2.3 Cartographie de flux (supérieurs à un critère ($F_{ij} > \alpha$)).....	9
2.3.1 Choix du critère de sélection des valeurs de flux	9
2.3.2 Choix du paramètre de taille graphique des flux (largeur du figuré)	9
2.3.3 Tracé des flux conformément aux deux critères (graphiques et thématiques)	10
2.3.4 Application d'une variable visuelle de Transparence aux figurés.....	11
2.3.5 Tracé de flèches.....	11
3. Procédure de sélection des flux à cartographier, selon un critère spatial (de voisinage binaire, ordinal ou de distance parcourue)	12
3.1 Préparation.....	12
3.1.1 Chargement des données numériques et spatiales (zones)	12
3.1.2 Préparation d'un fichier de points (x,y).....	12
3.2 Construction d'une matrice de voisinage ordinaire ($C_{ij}=1$)	13
3.3 Construction d'une matrice de voisinage ordinal (CK_{ij}).....	14
3.3.1 Construction de la matrice (CK_{ij}) avec $k=2$	14
3.3.2 Construction de la matrice (CK_{ij}) avec $k=4$	15
3.4 Construction et représentation des <i>k plus proches voisins</i>	17
3.4.1 Création du graphe de voisinage (k voisins)	17
3.4.2 Visualisation des graphes de voisinages.....	18
3.5 Construction d'une matrice de distances (D_{ij}) et réduction de (F_{ij}) en fonction de (D_{ij}) restreinte	19
3.5.1 Chargement des données numériques.....	19
3.5.2 Construction de la matrice de distance (D_{ij}) et réduction de (F_{ij})	20

4. Procédure de sélection des flux à cartographier, selon un critère global	22
4.1 Préparation.....	22
4.1.1 Chargement de la matrice (Fij) sous forme de liste.....	22
4.1.2 Préparation des matrices complémentaires – forme liste	22
4.1.3 Choix de la matrice à analyser (par exemple : Fij ⁺)	23
4.2 Analyse statistique de la normalité de (Fij).....	24
4.2.1 Histogramme de fréquences	24
4.2.2 Diagramme Quantile-Quantile	24
4.3 Analyse de la concentration des valeurs de flux.....	25
4.3.1 Calcul de l'indice de Gini.....	25
4.3.2 Courbe de concentration des valeurs de flux.....	26
4.4 Sélection globale selon un objectif de <i>significativité</i> de l'information.....	26
4.5 Sélection globale selon un objectif de <i>densité</i> de la matrice	27
5. Procédure de sélection des flux à cartographier, selon un critère local (« flux dominants »)..	28
5.1 Préparation.....	28
5.1.1 Connexion au répertoire de travail et chargement des packages requis.	28
5.1.2 Chargement de la matrice (Fij) à étudier	28
5.2 Sélection de flux majeurs et dominants (méthode classique Nuysten-Dacey, 1961)	28
5.2.1 Création et application d'une fonction de sélection des Flux majeurs	29
5.2.2 Création et application d'une fonction flux dominant.....	30
5.3 Sélection de flux dominants – variante sur le critère (1) : (i) envoie x% à (j) et (Dj)>(Di) ..	30
5.3.1 Création et application d'une fonction de calcul de fréquences	30
5.3.2 Application de la fonction flux dominant sur variante critère 1 exprimée en (α%)	31
5.3.3 Application de la fonction flux dominant sur variante critère 2.....	31
5.3.4 Variable de restriction de (Fluxtemp) en fonction de (fluxmajvar1).....	32
5.4 Sélection de flux dominants – variante sur le critère (2) : « ... et le volume de flux reçu par (j) est supérieur à celui reçu par (i) ».....	32
5.4.1 Variation du second critère sur (Fij) « ... et (Oi+Di) inférieur (Oj+Dj) »	32
5.4.2 Variable de restriction de fluxmajvar1 en fonction de flux_temp.....	33
5.5 Sous Programme de cartographie d'une des matrices de flux dominants	33
6. Procédure de cartographie directe d'une variable de l'environnement global de Rstudio©...	35
6.1 Préparation.....	35
6.2 Chargement de l'environnement global	35
6.2.1 Liste des variables de l'environnement	36
6.2.2 Cartographie d'une matrice située dans le "Global Environnement"	36
7. Références.....	38
8. Plan détaillé.....	39

