



**HAL**  
open science

## Simulation et validation de la dynamique urbaine par automate cellulaire

Edwige Dubos-Paillard, Patrice Langlois

► **To cite this version:**

Edwige Dubos-Paillard, Patrice Langlois. Simulation et validation de la dynamique urbaine par automate cellulaire. 2008. halshs-01694570

**HAL Id: halshs-01694570**

**<https://shs.hal.science/halshs-01694570>**

Preprint submitted on 27 Jan 2018

**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.

# Simulation et validation de la dynamique urbaine par automate cellulaire

## Simulation and validation of urban dynamics with cellular automata

DUBOS-PAILLARD Edwige, UMR CNRS 6049 THEMA, université de Franche-Comté, France  
[edwige.dubos-paillard@univ-fcomte.fr](mailto:edwige.dubos-paillard@univ-fcomte.fr)

LANGLOIS Patrice, UMR CNRS 6266 IDEES, université de Rouen, France  
[patrice.langlois@univ-rouen.fr](mailto:patrice.langlois@univ-rouen.fr)

### Résumé

Cet article présente les évolutions importantes la plateforme SpaCelle pour simuler l'évolution d'un espace à l'aide d'automate cellulaire. En particulier, la version 5 permet la structuration multi-couches des cellules pour gérer la dynamique de plusieurs variables définies sur un maillage. Elle permet aussi d'utiliser des cellules de type linéaires ou ponctuelles à travers des couches de type réseau. Les différentes couches ont leur propre dynamique et les règles peuvent faire interagir les couches entre elles. Ces fonctionnalités avancées ont permis de construire un nouveau modèle sur l'évolution de l'agglomération rouennaise plus proche de la réalité terrain en utilisant trois couches, dont une de type réseau. Le modèle a été validé par des techniques de comparaison de cartes relativement récente mieux adaptées aux problématiques de la croissance urbaine.

**Mots-clés** : simulation, modèle, automate cellulaire, croissance urbaine, validation.

### Abstract

This paper presents the important evolution of the SpaCelle platform to simulate the evolution of a space with cellular automata. In particular, the fifth version allows a multi layered structure of cells to manage several variable dynamics on a grid. It also permits the use of punctual and linear cells in network type layers. Each layer has its own dynamics and rules can make each layer interact. These advanced functions have allowed the construction of a new model of Rouen and its suburbs' evolution which is more realistic, using three layers of which one is a network type. The model has been validated with recent map comparison techniques which are more adapted to urban growth questions.

**Keywords**: simulation, model, cellular automata, urban growth, validation.

# 1 Introduction

Modéliser l'évolution urbaine c'est aussi s'interroger sur la « fabrique » de la ville au travers de l'expérimentation virtuelle. Néanmoins, il semble de plus en plus difficile de réduire la complexité de cette évolution à quelques modèles mathématiques préétablis, même si ceux-ci fournissent des pistes utiles d'explications, même partielles. Nous avons voulu ici, à travers l'outil SpaCelle, montrer qu'il est possible de développer un modèle qui n'est pas basé sur des équations mathématiques, mais sur des règles qualitatives d'évolution de l'espace, traduites assez directement du langage naturel. Ces règles s'appliquent en chaque lieu (les cellules de l'automate) et permettent de faire évoluer l'occupation du sol par la mise en concurrence entre les contraintes internes de développement de chaque cellule et des contraintes externes venues des cellules environnantes. Entre force de vie qui tente de faire perdurer l'état actuel et forces environnementales qui tentent de le changer, c'est toujours le plus fort qui gagne. Le système complexe – ville, ainsi modélisé, se transforme, dans un jeu permanent d'actions et de rétroactions. Nous montrons ici une modélisation de l'évolution de l'agglomération de Rouen (Haute-Normandie, France) durant la deuxième moitié du 20<sup>ième</sup> siècle qui a permis de valider ce genre de modélisation par rapport à une évolution connue. Cette agglomération occupe un vaste méandre de la Seine et compte environ 390 000 habitants en 1999. Néanmoins, l'espace étudié s'étend au-delà du continuum urbain et couvre un domaine de 345 km<sup>2</sup> incluant une grande partie des communes périurbaines.

Après avoir rappelé les fonctionnalités principales de SpaCelle, en particulier son langage simple et son organisation multicouches qui permettent de définir plusieurs populations cellulaires de types différents en interactions, nous présentons un nouveau modèle de dynamique urbaine sur Rouen qui est soumis à différentes méthodes de validation en partie inspirées des travaux du groupe RIKS (Hagen 2002) et (Hagen-Zanker A., Straatman B., Uijee I., 2005), basées sur la comparaison entre des cartes observées avec les cartes simulées.

## 2 La version 5 de SpaCelle, automate cellulaire multicouches

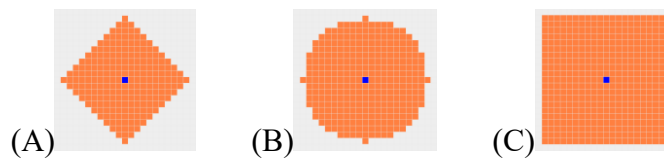
### 2.1 Présentation générale

Rappelons que SpaCelle, présenté dans sa version 3 dans (Dubos-Paillard & al. 2003) est une mini-plateforme à base d'automate cellulaire. SpaCelle (maintenant en version 5) est une application permettant à l'utilisateur de décrire entièrement le modèle dynamique par un ensemble de règles et de paramètres généraux (taille du domaine, mode d'exécution, etc.) accessibles par l'interface utilisateur (taille du domaine, mode d'exécution, etc.). Le langage de modélisation, propre à l'outil et donc plus simple que les langages de programmation, permet de décrire les règles de la dynamique, mais aussi les couches cellulaires, et les états du modèle. L'utilisateur doit aussi alimenter le modèle en entrée par une configuration initiale des états, par saisie directe sur l'écran ou importation de données. Il peut aussi paramétrer la forme et la taille du domaine, la géométrie des cellules (carré ou hexagone pour un maillage), le mode de fonctionnement global du modèle (synchrone, asynchrone, ou aléatoire). Il doit aussi choisir la métrique (topologie) et la fonction de pondération de l'interaction en fonction de la distance applicable à son modèle. Il peut alors procéder à des simulations et observer certaines sorties du modèle dynamique en cours d'exécution (listes d'images, graphiques, etc.).

La version 5 du logiciel est multi-couches, un modèle se compose alors d'une ou plusieurs couches cellulaires et d'une base de règles. Le modèle est stocké dans un fichier selon un format texte propriétaire (extension ac5). Une couche est composée soit d'un maillage régulier de carrés ou d'hexagones, soit d'un réseau linéaire composé de points ou de lignes polygonales. Les couches peuvent être importées directement d'une base de données géographique (format GRID ASCII pour les couches cellulaires et format vectoriel pour les couches réseau). Les couches ainsi importées servent de configuration initiale pour l'automate cellulaire. Pour faire une simulation, l'utilisateur peut soit charger un modèle existant depuis un fichier au format SpaCelle, soit définir de toute pièce un nouveau modèle par saisie ou importation de données externes.

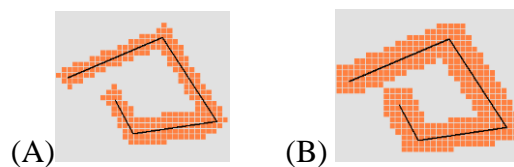
## 2.2 Le choix de la métrique et la notion de voisinage dans SpaCelle

L'utilisateur choisit l'une des trois métriques  $d_1$  (distance Manhattan),  $d_2$  (métrique euclidienne) et  $d_\infty$  (distance du max) pour définir la topologie de voisinage de son modèle qui s'applique aussi bien pour les couches maillées que pour les couches réseau. Ce n'est donc pas une topologie de contiguïté entre objets mais une topologie de proximité. Ainsi, dans SpaCelle, Le voisinage de rayon  $R$  relatif à une cellule  $C$  est l'ensemble des cellules situées à une distance inférieure ou égale à  $R$  de la cellule  $C$ . La forme d'un voisinage dépend donc de la métrique choisie, comme on peut le voir sur les figures 1 et 2 suivantes.



**figure 1.** voisinages de rayon  $R = 10$  sur un maillage carré

(A): distance de Manhattan, (B): distance euclidienne, (C): distance du Max



**figure 2.** voisinages d'une cellule linéaire dans une couche cellulaire.

(A) : distance de manhattan, (B) : distance du max

## 2.3 Le comportement cellulaire sous SpaCelle

Pour modéliser un phénomène géographique, le mécanisme de transition cellulaire peut être complexe. Pour le définir avec un maximum de souplesse, on l'exprime par un ensemble de règles indépendantes réparties en deux catégories : les **règles de vie** (événements internes) et les **règles de transition** (événements externes).

La syntaxe générale de toute règle est de la forme suivante :

EtatInitial > EtatFinal = Pertinence (formule 2)

La partie « EtatInitial > EtatFinal » est la partie **action** de la règle, le changement d'état ou transition. Le symbole « = » sépare simplement les deux parties de la règle, sans autre signification. La partie « Pertinence » permet d'évaluer la pertinence (la force) de la transition.

Elle se présente sous la forme d'une expression dont l'évaluation pour une cellule donnée, à un instant donné, est un nombre réel entre 0 et 1.

## 2.4 Structure générale d'un modèle

Enfin, il doit modifier ou saisir les règles de transition de son modèle. Ces règles constituent la base de connaissances de la dynamique du modèle.

Les cellules d'une même couche sont caractérisées par un ensemble d'états possibles. A chaque état est associée une représentation symbolique qui sert à la visualisation. A chaque couche cellulaire correspond un mode de représentation. Par défaut, le logiciel remplit la cellule d'une couleur associée à son état, mais on peut aussi – en particulier lorsqu'il y a plusieurs couches de cellules à superposer – associer un type de symbole à la couche, défini par sa forme, la couleur et l'épaisseur du contour du symbole. A chaque état de cette couche on associe une couleur de remplissage du symbole ou de la cellule. Le domaine peut être recouvert par plusieurs couches cellulaires dont chacune peut posséder sa propre dynamique ou être statique. Chaque couche peut influencer la dynamique des autres couches, car une même règle peut combiner des informations relatives à plusieurs couches.

## 2.5 Le comportement cellulaire, naissance, vie et mort, loi de la règle la plus pertinente

Une cellule dans un état donné peut être considérée comme un agent qui possède une certaine durée vie (définie par une règle de vie propre à cet état). L'âge de l'agent conditionne sa « force de vie » qui vaut 1 à sa naissance, et décroît linéairement jusqu'à 0 à sa mort naturelle. Comme cette force diminue au cours du temps, l'agent subit de plus en plus la concurrence des « forces environnementales » exercées par les agents voisins. Si, à un instant donné, une force environnementale dépasse sa propre force de vie, l'agent « meurt », c'est-à-dire subit une transition vers l'état défini pour sa mort dans sa règle de vie.

Lorsqu'une cellule prend un nouvel état, cela correspond à la naissance d'un agent cellulaire qui est alors affecté d'une durée maximale de vie (dans cet état) qui dépend de sa classe (DI : durée infinie, DF : durée fixée, DA : durée aléatoire selon une espérance de vie et un écart-type). Si l'agent atteint l'âge de sa mort naturelle, il change d'état et prend l'état de mort défini dans la règle de vie de sa classe.

Le comportement cellulaire est basé sur un *principe décisionnel unique*, parmi les règles applicables à un instant donné pour une cellule, *c'est la règle la plus pertinente qui s'applique toujours*. Chaque règle qui s'applique à un agent cellulaire calcule une valeur qui s'interprète comme une *force* ou une *pertinence* d'application de la règle. C'est celle dont la valeur est maximale qui est exécutée. Si plusieurs règles atteignent le maximum, une seule règle est retenue parmi celles-ci par tirage aléatoire. En d'autres termes, ce principe choisit la règle associée à la plus grande valeur entre la « force de vie » d'une cellule et les « forces environnementales » émanant des autres cellules. Ce choix décide à chaque instant, quelle action de transition devra s'appliquer à la cellule.

On voit que ce mécanisme est très différent d'une description algorithmique du comportement cellulaire ou selon des règles d'inférences logiques. La base de règle n'est pas une liste d'instructions du type « si *condition* alors *action* ». A chaque instant et pour chaque cellule, une seule règle d'action (la plus pertinente) est choisie et exécutée. Ainsi, l'ordre des

règles n'a aucune importance, certaines peuvent être inutiles, elles peuvent même être contradictoires entre elles, le jeu de la concurrence fera toujours son choix.

## 2.6 Fonctionnement déterministe ou stochastique

On voit donc que le mécanisme de décision est généralement déterministe, sauf en cas d'égalité entre les règles les plus pertinentes. D'autre part les durées de vie des agents cellulaires peuvent être déterministes ou aléatoires. Il existe aussi des fonctions spéciales qui déclenchent des événements, soit à un instant tiré aléatoirement, soit à une date définie. Ces fonctions événementielles, par combinaison avec d'autres fonctions d'évaluation, permettent d'introduire ou non une dose d'aléatoire. Enfin, même si les cellules possèdent une durée de vie fixe, leur âge à l'instant initial est défini de manière aléatoire par le programme.

On voit donc que le modélisateur peut faire son choix entre un modèle complètement déterministe (comme le jeu de la vie), fortement stochastique (comme le modèle de discrimination socio-spatial de Schelling par exemple), ou ayant à la fois une composante déterministe et une certaine dose d'aléatoire.

## 2.7 Les règles

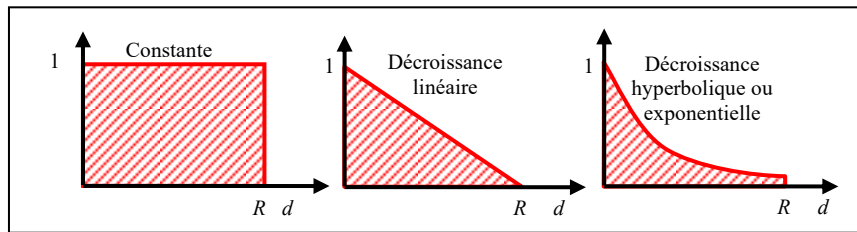
Nous ne développerons pas ici la manière d'écrire les règles, déjà décrite dans (Dubos-Paillard & al. 2003), nous donnons juste un exemple pour rappeler la syntaxe de ce langage. Ainsi la phrase : « une friche peut devenir pavillonnaire s'il existe déjà du pavillonnaire ou du commerce dans le voisinage et s'il n'y a pas d'industrie dans un voisinage un peu plus éloigné » se traduit par la règle de transition suivante :

$$\text{Fri} > \text{Pav} = \text{EV}(\text{Pav} + \text{Com} ; 3) * \text{ZV}(\text{Ind} ; 5)$$

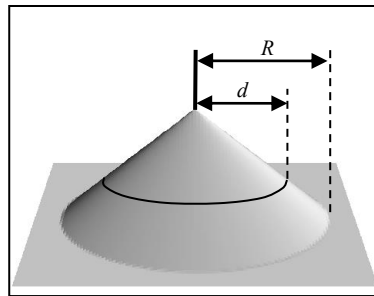
Dans laquelle les notations « Fri », « Pav », « Com », « Ind » sont des symboles d'état de la couche d'occupation du sol associés à « friche », « habitat pavillonnaire », « commerce » et « industrie ». La conjonction « et » de la phrase se traduit par une multiplication notée « \* », alors qu'un « ou » se traduit par une addition « + ». La fonction EV signifie « il Existe dans le Voisinage » et ZV signifie « il n'existe pas dans le voisinage » (ZV comme Zéro dans le Voisinage). Il existe un quinzaine de fonctions de ce genre. Les valeurs 3 et 5 représentent les rayons des deux voisinages envisagés. On peut expérimenter plusieurs valeurs de rayon, avant de fixer celle qui convient le mieux.

## 2.8 Evaluation stricte ou floue d'un voisinage

Les fonctions comme PV, « Proportion dans le Voisinage » qui sont de la forme  $F(X, R)$  évaluent leur voisinage soit en considérant tous les voisins de la même manière (pondération constante), comme dans la figure 3a-(1), soit en utilisant une pondération décroissante en fonction de la distance (figure 3(2) ou 3(3)). L'utilisateur a le choix entre quatre pondérations : constante  $g(d) = 1$ , linéairement décroissante  $g(d) = 1-d/R$ , décroissance hyperbolique  $g(d) = 1/(d+1)$  ou plus généralement en décroissance exponentielle  $g(d) = (d+1)^{-k}$ .



**figure 3.** Les fonctions de pondération



**figure 4.** Evaluation d'un voisinage pondéré linéairement

L'évaluation de  $F(X, R)$  sur un voisinage de rayon  $R$ , est définie par une fonction  $f$  qui calcule la densité d'agents de type  $X$  sur le cercle de rayon  $d$  (fig. 4). Le résultat  $F(X, R)$  est la moyenne de ces densités pondérée par  $g$ :

$$F(X, R) = \frac{\sum_{d=0}^R g(d)f(d, X)}{\sum_{d=0}^R g(d)}$$

## 2.9 Les sorties

Le logiciel permet de sélectionner un certain nombre de sorties, lui permettant d'observer le comportement du modèle en cours ou après la simulation :

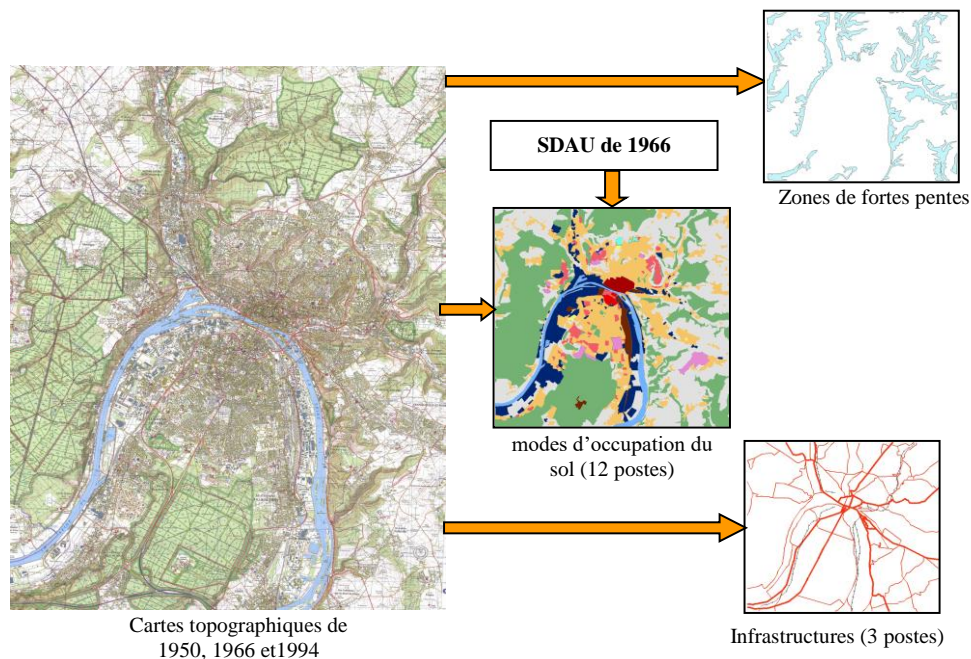
- Cartes des configurations successives, avec enregistrement possible de la suite d'images (permettant de faire le film de la simulation).
- Observation globale de la dynamique par des graphiques d'évolution de la fréquence des états.
- Observation locale qui visualise les règles applicables à chaque cellule et leur pertinence.
- Observation de la fréquence d'utilisation des différentes règles au cours d'une simulation, ce qui permet d'analyser le rôle et l'importance de chaque règle au cours de la simulation.
- On peut aussi enregistrer une configuration finale ou intermédiaire en format « grid », ce qui permet de réutiliser les résultats dans un SIG, ou dans un logiciel de comparaison de carte comme « Map Comparison Kit » (Hagen-Zanker et al).

## 3 Le modèle d'évolution de l'agglomération rouennaise

Après une première modélisation (Dubos-Paillard et al, 2003) de l'espace rouennais, un nouveau modèle a été réalisé, basé sur les possibilités multi-couches du logiciel SpaCelle (version 5) pour simuler la croissance urbaine de l'agglomération rouennaise entre 1950 et 1994. La philosophie du modèle reste la même mais les règles de transition ont connu d'importantes transformations dans la mesure où dans la version 3 chaque cellule, de type surfacique, ne pouvait contenir qu'un état. Ainsi, dans le premier modèle, une route était représentée par une juxtaposition de cellules entières, ce qui prenait une surface considérable par rapport à la réalité (chaque cellule couvrant 2,25 ha) et empêchait de représenter le type d'occupation du sol sur ces surfaces. De même les zones de forte pente ne pouvaient avoir que l'état « forte pente ». Il était assez surprenant qu'il puisse être remplacé par un état « habitat » car ces deux types d'informations n'étaient pas exclusives l'une de l'autre.

### 3.1 De la carte à l'automate cellulaire

Les modes d'occupation de l'espace observé ont été définis à partir de cartes topographiques disponibles à différentes dates (1950, 1966 et 1994). Les informations obtenues ont pu être affinées grâce à une carte issue du SDAU de 1966 présentant précisément la localisation des différentes fonctions dans l'espace étudié.



*figure 5. Construction des couches de l'automate*

Pour chaque année observée, nous avons extrait 3 couches d'informations comportant respectivement les modes d'occupation du sol, les secteurs à fortes pentes et les principaux axes de transport qui ont été isolés dans une couche de type linéaire (figure 5). Ces informations géographiques ont permis de définir les états des cellules des différentes couches de l'automate. Dans notre modèle, seule la couche des pentes reste statique, les deux autres, celle des axes de transport et celle comportant les modes d'occupation de l'espace possèdent une dynamique gérée par le modèle. L'espace cellulaire se présente donc sous la forme d'une



grille de 121 lignes et 127 colonnes constituée de cellules carrées de 150m de côté (soit 2,25 hectares), soit un total de 15367 cellules (soit 345,75 km<sup>2</sup>).

## 3.2 La base de règles

### 3.2.1 La question de la durée d'activité des règles

La base de règles a été construite à partir d'une analyse des grandes étapes de l'urbanisation depuis les années 1950. Cette période a été marquée par une diversification des modes d'occupation du sol et l'apparition de nouvelles logiques spatiales modifiant la morphologie des villes françaises telles que le zonage des espaces urbains, la périurbanisation, l'exurbanisation ou le desserrement de certaines catégories d'activités économiques qui privilégient des zones d'activités plus moins spécialisées ou de vastes zones commerciales à l'interface de l'agglomération et de l'espace périurbain. Ces nouvelles formes d'occupation de l'espace relèvent de deux catégories. La première se compose d'espaces construits au cours d'une période facilement identifiable. Ainsi, la construction des grands ensembles en France débute en 1953-1954, avec notamment l'action de l'abbé Pierre et se termine au début des années 1970, la circulaire "Guichard" de 1973 portant un coup d'arrêt à leur construction. La seconde catégorie concerne des processus tels que la périurbanisation qui ne se traduisent pas forcément par une modalité spécifique d'occupation de l'espace. Le lotissement est certes une traduction spatiale de ce processus mais il n'est pas le seul. De ce fait, il est difficile d'établir une date de commencement pour la simulation même si l'on admet que le processus débute dans les années 60 autour des grandes villes et se développe à partir des années 1970. Nous avons donc considéré, un peu arbitrairement, que l'année 1970 correspondait dans notre modèle au début du processus et qu'il était toujours à l'œuvre même si certains chercheurs évoquent une réduction de son ampleur durant les années 1990.

### 3.2.2 Les règles du modèle

Le modèle contient 16 types d'agents-cellules dont les durées vie peuvent être, soit infinies (fleuve, forêt, grands équipements, etc), soit aléatoires (espaces non bâtis, habitat individuel, zones industrielles, etc). Elle contient aussi une trentaine de règles de transition d'importance inégale qui n'agissent pas nécessairement sur l'ensemble de la période. Ainsi, les règles concernant les grands ensembles deviennent inactives à compter de 1972. En revanche, les règles sur l'apparition des centres commerciaux, la création des parcs tertiaires se déclenchent respectivement au milieu des années 1960 et au milieu des années 1980. D'une manière générale, les règles utilisées se divisent en 3 catégories:

- Des règles principales d'apparition des formes d'occupation du sol :

Ce sont des règles relativement complexes qui contiennent l'essentiel des connaissances et des hypothèses sur les mécanismes de la croissance urbaine. Elles représentent 40% des règles de la base.

**Exemple 1 :**

$$F_{ri} > H_{pd} = PV(H_{pd} + Ch + Hd + Eqp + Imm; 10; 0.58; 1) * AV(Z_i; 1) * ZV(P_{en}; 2) * ZV(A_{fe}; 1) * AV(Se_i + V_{fe}; 2) * PV(Imm; 1; 0; 0.25) * PV(Z_i + V_{fe}, 6; 0; 0.1)$$

L'exemple 1 peut être interprété de la façon suivante : l'espace non bâti ( $F_{ri}$ ) devient de l'habitat peu dense ( $H_{pd}$ ) si le voisinage de la cellule, de rayon 1,5km, est déjà sensiblement urbanisé, s'il n'y a pas à proximité immédiate de zone industrielle ( $Z_i$ ), de relief ( $P_{en}$ ), de voie ferrée ( $A_{fe} + V_{fe}$ ) et la Seine ( $Se_i$ ). Il faut également que la proportion de grands ensembles

(Imm) ne soit pas trop grande au sein des cellules voisines et que la densité d'industries (zi) soit faible dans un rayon de 900m.

- Des règles de cohérence spatiale

Ces règles ne sont pas des règles contenant des connaissances ou des hypothèses concernant le développement de l'espace étudié. Ces règles de cohérence spatiale sont destinées à éviter des situations rarement observées dans la réalité comme le fait d'avoir une industrie au milieu d'un quartier de grands ensembles ou d'avoir un centre commercial ou un quartier de grands ensembles de forme effilée. Elles représentent 44% des règles et perdurent sur l'ensemble de la période pour la plupart.

**Exemple 2 : Fri>Zi = PV(Zi;1;0.6;1)**

Ces règles peuvent être simples ou plus complexes. L'exemple 2 indique que l'espace non bâti devient industriel si le voisinage de la cellule est composé en forte proportion d'industries. Il est possible d'ajouter des contraintes supplémentaires dans ce type de règles. Elles peuvent porter sur un nombre maximal de cellules relevant d'un état donné dans le domaine, sur la présence ou l'absence d'un état spécifique dans un rayon défini etc.

- Des règles d'aléa « encadré »

Ces règles concourent à la dimension stochastique du modèle. Dans le modèle sur Rouen, elles représentent 16% des règles de transition mais leur impact sur le résultat final reste limité.

**Exemple 3 : Fri>Hpd = AL(20;1) \* EV(Hpd+Imm+Ch+Hd;1)**

L'exemple 3 indique que, de façon aléatoire, chaque année, 20 cellules d'état friche deviennent de l'habitat à condition qu'il existe de l'habitat à proximité.

### **3.3 Evolution observée de l'espace rouennais entre 1950 et 1994**

La comparaison de l'occupation du sol entre 1950 et 1994 montre que près de 18% des cellules ont changé d'état en 45 ans (figure 6). L'analyse de la carte de 1966 permet de distinguer deux périodes notablement différentes. Entre 1950 et 1966, à peine 5% des cellules changent d'état du fait d'une faible extension du bâti qui se fait principalement sur les périphéries de l'espace aggloméré. Elle prend surtout la forme de grands ensembles qui donnent lieu à de nouveaux quartiers localisés fréquemment sur des terrains bon marché (ex Canteleu) et d'un habitat individuel qui s'étend principalement vers le sud de l'agglomération. Les industries apparaissent principalement en aval de Rouen, dans la zone industrialo-portuaire.

La période 1966-1994 connaît davantage de modifications puisque 15,4% des cellules changent d'état. Celles-ci touchent en premier lieu l'habitat individuel qui continue à s'étendre sur les pourtours de l'espace aggloméré mais aussi, à partir des années 1970, autour des villages ruraux (périurbanisation). Les grands ensembles connaissent une croissance limitée puisqu'au début des années 1970, ils sont quasiment tous achevés. L'industrie progresse également. La zone industrialo-portuaire, en aval de Rouen continue à se développer vers le sud. Au cours des années 1970 et une vaste zone d'activité apparaît en amont de Rouen, en bordure de Seine.

C'est durant cette période que se développent les centres commerciaux et des zones d'activités dédiées principalement aux fonctions tertiaires qui se localisent au nord de l'agglomération. Enfin le réseau routier se développe notablement après 1966.

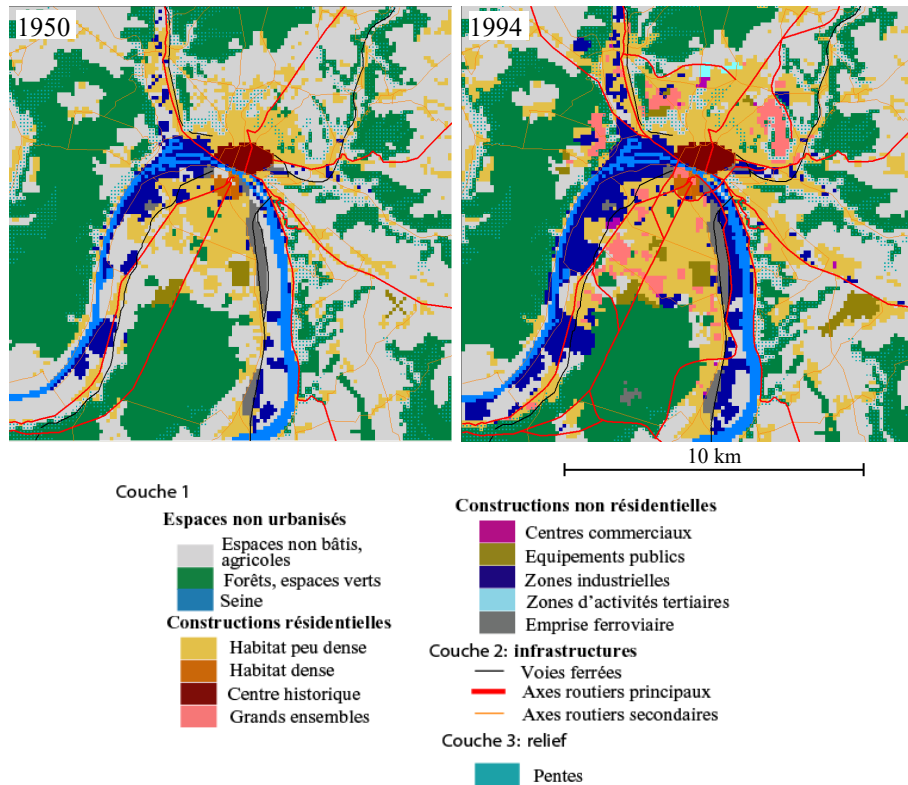


Figure 6. Simulation observée en 1950 et 1994

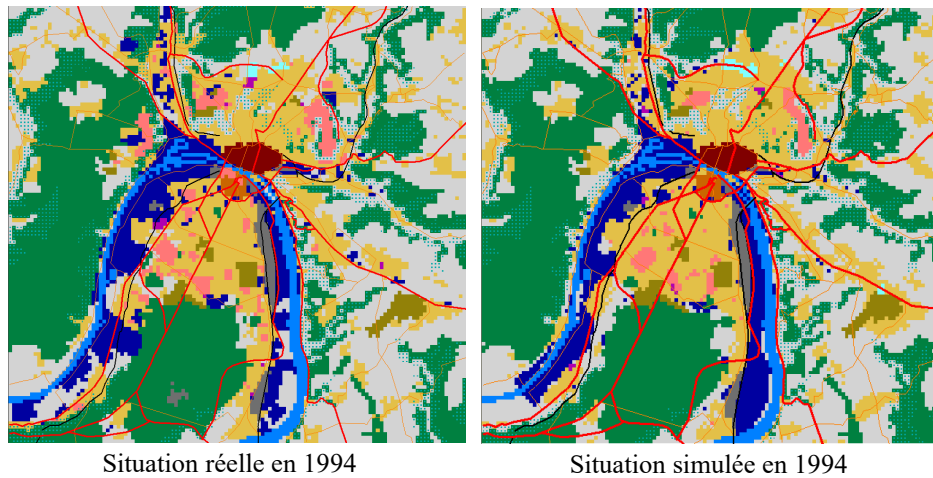
### 3.4 Résultats et validations de la simulation entre 1950 et 1994

L'état des lieux de l'évolution de l'espace rouennais a permis d'alimenter les règles de notre modèle. Elles définissent des contraintes spatiales fortes laissant néanmoins certains degrés de liberté. Cette liberté s'exprime à travers une composante stochastique qui permet d'obtenir des résultats légèrement différents à chaque simulation. Nous présentons donc ici un exemple de simulation entre 1950 et 1994.

#### 3.4.1 Comparaison visuelle en fin de simulation (1994)

La figure 7 compare les situations observées et simulées en 1994. D'un point de vue global, on constate que la situation simulée est assez proche de la situation observée. Dans le détail, on remarque que dans la situation simulée, l'habitat s'est un peu plus développé vers le nord et l'est. Cela s'explique en partie par le fait que la proximité des principaux axes routiers a été considérée comme un facteur important du développement de l'habitat dans les espaces périurbains. On peut aussi constater que dans la situation simulée l'industrie s'est plus développée en amont de Rouen et en aval dans le secteur de Petit-Couronne et Grand-Couronne au sud-ouest de l'agglomération. Enfin, on note que quelques configurations spatiales n'émergent pas de la simulation, même en faisant varier légèrement les contraintes des différentes règles de transition. L'exemple le plus significatif est celui des grands ensembles localisés sur le plateau situé à l'ouest du centre-ville (Canteleu). Ces grands ensembles ont été construits sur un espace entouré de forêts, mal relié au centre de

l'agglomération, en retrait des axes routiers principaux et de l'habitat individuel. Il s'avère que l'explication de cette implantation est spatialement illogique. Elle tient surtout à une volonté politique et au coût très faible des terrains qui ont pu être acquis pour seulement 0,15 euro le mètre carré.



*figure 7. Comparaison entre situation observée et simulée en 1994*

### 3.4.2 Validation: L'apport des méthodes basées sur les ensembles flous

La validation des simulations est une étape importante non seulement pour confirmer le modèle retenu mais également pour l'améliorer. Plusieurs méthodes de comparaison cellule par cellule sont couramment utilisées. Elles sont basées sur le taux de similarité qui rapporte le nombre de cellules identiques aux deux cartes au nombre total de cellules. Mais face aux biais importants associés à ces méthodes et considérant que la validation par comparaison visuelle a toute sa légitimité, certains chercheurs ont proposé d'autres méthodes tenant compte des propriétés du voisinage des cellules. « The strong side of fuzzy map comparisons is that it resembles the way human observers compare maps » (Visser, H., & de Nijs, T<sup>1</sup>). Ces dernières années, le laboratoire RIKS<sup>2</sup> a développé un outil, Map Comparison Kit, qui propose, outre les méthodes de validation classiques, des méthodes intéressantes dans la mesure où elles tiennent compte des propriétés du voisinage de chaque cellule. Nous présentons ici les résultats de deux d'entre elles.

La première est un test non paramétrique défini par Cohen (Cohen 1960), le coefficient Kappa ( $K$ ) qui est basé sur la comparaison de deux observations (nos deux cartes) d'une même série de  $n$  objets (ici nos cellules) pouvant prendre  $r$  modalités qualitatives (nos classes d'occupation du sol). Le coefficient Kappa exprime la différence relative entre la proportion de concordance observée  $P_o$  et la proportion de concordance aléatoire  $P_e$  relativement à la différence entre la concordance parfaite et aléatoire.  $K$  exprime donc le pourcentage d'écart à la concordance maximale corrigé du simple effet du hasard. Ainsi, le résultat est une valeur comprise entre -1 (discordance totale, ), 0 (choix aléatoire) et 1 (entièrement concordant).

$$K = \frac{P_o - P_e}{1 - P_e}$$

Pour calculer la concordance observée  $P_o$  et aléatoire  $P_e$ , on comptabilise dans un tableau de contingence les effectifs croisés des différentes modalités des cellules des deux cartes. Si

<sup>1</sup> [http://storage.riks.nl/pub/MCK/MCK\\_Reader.pdf](http://storage.riks.nl/pub/MCK/MCK_Reader.pdf)

<sup>2</sup> Research Institute for Knowledge Systems, voir le site: <http://www.riks.nl/>

l'on note  $n_{ij}$  le nombre de cellules qui possèdent la modalité  $i$  dans la première carte et  $j$  dans la deuxième carte, la concordance observée est la proportion de cellules qui possèdent la même modalité dans les deux cartes, c'est-à-dire la proportion des cellules classées dans les cases diagonales du tableau de contingence, soit :

$$P_o = \frac{1}{n} \sum_{i=1}^n n_{ii}$$

Si les cartes étaient construites au hasard en respectant leurs effectifs pour chacune des  $r$  modalités, qui sont les sommes marginales  $n_{i\bullet}$  et  $n_{\bullet i}$  du tableau de contingence, les cartes seraient alors statistiquement indépendantes et l'effectif théorique serait défini par  $t_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}$ . La concordance aléatoire  $P_e$  est la proportion d'effectifs diagonaux de cette matrice de contingence théorique (qui respecte l'indépendance). Ainsi :

$$P_e = \frac{1}{n} \sum_{i=1}^r t_{ii} = \frac{1}{n^2} \sum_{i=1}^r n_{i\bullet} \cdot n_{\bullet i}$$

La seconde méthode, dite « coefficient Kappa flou » (ou « fuzzy kappa »),  $K_F$ , tente de tester la concordance en se rapprochant de l'impression visuelle. Elle se base sur le principe qu'il n'est pas nécessaire de faire correspondre pixel à pixel les cartes observée et simulée, mais recherche s'il existe une forte similitude de voisinage entre les cartes en présence. Un exemple est donné ci-dessous. Les deux figures (8A et 8B) comptent un nombre identique de cellules grises et noires, placées en alternance dans les deux cas. Visuellement, nous avons deux situations proches qui donnent une discordance totale en utilisant le coefficient de Kappa qui indique des situations entièrement différentes puisque la première case est noire dans la figure 8A et grise dans la figure 8B.

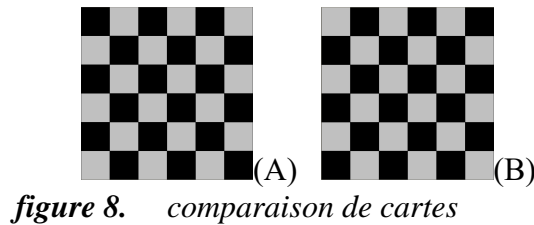


figure 8. comparaison de cartes

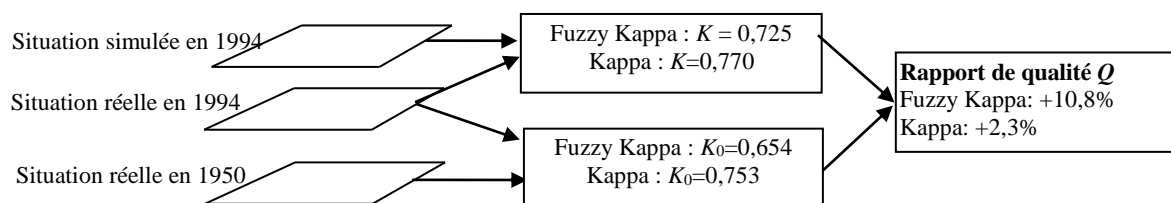
Le « Kappa flou » s'écrit encore sous la même forme que  $K$  :

$$K_F = \frac{P_{Fo} - P_{Fe}}{1 - P_{Fe}}$$

mais introduit deux assouplissements dans le calcul : les proportions floues observée  $P_{Fo}$  et aléatoire  $P_{Fe}$  sont calculées en tenant compte à la fois d'un certain écart de localisation entre les pixels comparés (cet écart est quantifié par une fonction décroissante de la distance dans un voisinage) et d'une certaine ressemblance modale (par une matrice de similarité entre les  $r$  modalités de la variable descriptive). La proportion d'accord de la carte A par rapport à B est calculée en mesurant pour chaque pixel  $x_i$  de A, quel est le pixel  $y_j$  de B au voisinage de  $i$  qui réalise le meilleur compromis entre distance et ressemblance avec  $x_i$ . On calcule ensuite la moyenne de ces mesures pour tous les pixels et on fait de même avec la carte B par rapport à A. On garde comme valeur finale pour établir les degrés de similitude  $P_{Fo}$ , la valeur minimale obtenue par les deux sens de calcul. Le calcul de  $P_{Fe}$  correspond à l'espérance mathématique

de cette même mesure calculée à partir de cartes *A* et *B* tirées au hasard tout en respectant les fréquences des modalités observées de *A* et *B*, c'est-à-dire deux cartes idéalement indépendantes.

Notons cependant que, quelle que soit la méthode retenue, les résultats doivent être analysés avec précautions dans la mesure où ils restent liés au nombre d'états et à l'importance des changements intervenus au cours de la période d'observation même si les coefficients du Kappa et du Kappa flou ambitionnent d'éliminer la part de concordance due à la chance. C'est la raison pour laquelle Alex Hagen-Zanker propose d'utiliser un rapport de qualité plus exigeant, établi à partir de cartes historiques et prenant davantage en compte les effets de structure. Comme l'indique la figure ci-dessous, il s'agit d'un indicateur qui compare les résultats obtenus  $K$  à une valeur de référence  $K_0$  qui est pour le cas présent établie en appliquant les mêmes méthodes sur la comparaison des situations réelles en début et en fin de période. Un ratio de type  $Q = \frac{K - K_0}{K_0}$  est ainsi calculé entre les deux valeurs obtenues. Une valeur positive indique un résultat correct tandis qu'une valeur négative indique un résultat qui est d'autant moins bon que la valeur est grande.



D'après A.Hagen-Zanker

**figure 9.** Relativisation des résultats obtenus à l'aide des méthodes Kappa et Fuzzy Kappa

L'application sur le cas rouennais donne un résultat pour le coefficient de Kappa qui est correct sur la période 1950-1994 (figure 9).

$K = \text{simulé 1994} / \text{réel 1994}$	0.770
$K_0 = \text{réel 1994} / \text{réel 1950}$	0.753
rapport de qualité Q	+2.3%

**Tableau 1 :** Comparaison entre situation observée et situation simulée à l'aide du coefficient de Kappa

L'utilisation de la méthode « fuzzy kappa » donne des résultats différents selon la fonction décroissante, le rayon et le paramètre retenus. Néanmoins tous vont dans le même sens comme le montre le tableau 2 qui présente différents résultats obtenus en utilisant une fonction décroissante exponentielle et différentes valeurs pour le rayon et le paramètre.

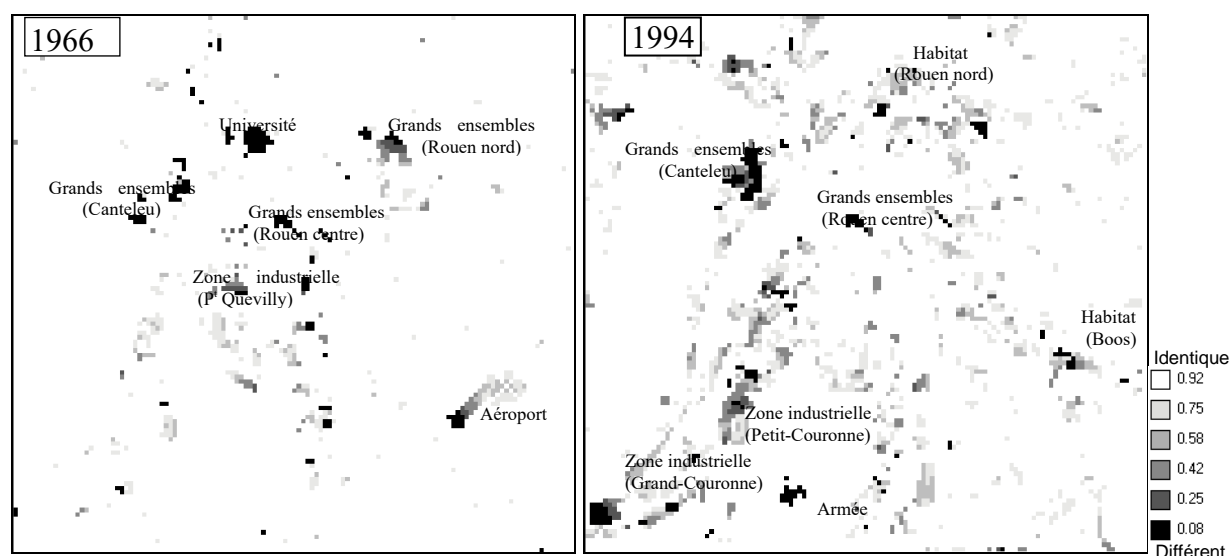
		K= simulé 1994 / réel 1994			K <sub>0</sub> = réel 1994 / réel 1950			rapport de qualité		
		Demi effet de la distance						Q = (K-K <sub>0</sub> )/K <sub>0</sub>		
			1	2	4	1	2	4	1	2
Rayon	2	0.751	0.713	0.682	0.7	0.652	0.598	7.3%	9.4%	14.0%
	4	0.744	0.723	0.696	0.701	0.651	0.582	6.1%	11.1%	19.6%
	8	0.745	0.725	0.702	0.702	0.654	0.589	6.1%	10.9%	19.2%
	10	0.745	0.725	0.703	0.702	0.655	0.591	6.1%	10.7%	19.0%
	20	0.745	0.725	0.705	0.702	0.655	0.593	6.1%	10.7%	18.9%

**Tableau 2 :** Comparaison entre situation observée et situation simulée à l'aide du coefficient de Kappa

Si le calcul du kappa flou donne des résultats intéressants, une étude plus fine tenant compte des différents modes d'occupation du sol montre d'une part qu'il est plus facile d'avoir de bons résultats quand l'effectif de l'état est important et d'autre part que le rapport de qualité ne peut être établi lorsque l'état n'est pas présent en début de période.

Pour les représentations cartographiques (figures 10) nous avons retenu un rayon de 8 cellules (1200 mètres) avec un demi-effet à 2 cellules (300m). Il ressort en 1966 que les principaux écarts entre réel et simulé sont dus soit à la création ou l'extension de grands équipements (aéroport au sud-est, université au nord) soit à la réalisation de grands ensembles qui n'apparaissent pas dans la simulation. La mise en évidence des grands équipements est logique puisque nous avons pris le parti de ne pas simuler leur développement et de les implanter au début de chaque période analysée c'est-à-dire en 1950 et en 1966. Concernant les grands ensembles les situations sont variables. Pour Rouen-nord, il s'agit seulement d'un décalage temporel. En revanche, pour Canteleu et Rouen-centre le modèle utilisé n'est pas en mesure de les faire apparaître d'où leur mise en évidence sur les cartes de 1966 et 1994.

En 1994, outre les grands ensembles évoqués ci-dessus, les conclusions visuelles concernant l'habitat et l'industrie sont confirmées. La carte des écarts à la réalité met davantage en évidence la zone industrielle en aval de Rouen au niveau de Petit Couronne et Grand-Couronne. Concernant l'habitat, nous pouvons également noter que le modèle n'a pas tenu compte de l'aéroport dans le développement des espaces périurbains. De ce fait, les constructions dans le secteur de Boos (Figures 10) sont plus importantes que dans la réalité. Ces différents constats sur l'industrie et l'habitat sur la période 1966-1994 doivent nous conduire à réajuster les règles concernées sans les modifier fondamentalement. Il faudra cependant intégrer une contrainte supplémentaire limitant l'expansion de l'habitat à proximité de l'aéroport. Ceci peut être fait à l'aide d'une fonction de type AV (absence dans le voisinage) ou PV (proportion dans le voisinage).



*figure 10. Comparaison entre situation observée et situation simulée à l'aide de la méthode fuzzy Kappa en 1966 et 1994*

## 4 Conclusion

Une première conclusion concerne l'apport de cette nouvelle modélisation par rapport au modèle développé dans (Dubos-Paillard et al. 2003). La grande supériorité de ce dernier provient des possibilités de séparation du réseau et des surfaces. En séparant les informations en plusieurs couches, surfaciques, linéaires ou ponctuelles, on peut modéliser la dynamique de chaque couche mais aussi leur interaction. Ainsi, le réseau se développe en fonction de l'occupation du sol, mais le réseau stimule en retour ce développement. Par contre, des couches plus statiques comme le relief, permettent de moduler efficacement la dynamique de développement sans gêner la couche ayant la dynamique principale.

La deuxième conclusion concerne la validation. Les possibilités offertes pour la comparaison d'un espace observé et d'un espace simulé ont pour but premier la validation du modèle retenu. Même si cet article n'était pas centré sur ce sujet, nous avons néanmoins réalisé des comparaisons de cartes pour valider le modèle mais aussi afin de l'améliorer pour des simulations futures.

Enfin, l'ambition du travail effectué jusqu'à présent a consisté à simuler la croissance urbaine de l'agglomération rouennaise sur une longue période (1950-1994) afin de cerner précisément les mécanismes spatiaux en vigueur et montrer la pertinence de ce type de modélisation. L'intérêt de travail est que les résultats issus de la simulation peuvent être confrontés à d'autres méthodes d'analyse de la croissance urbaine (par exemple plus orientées vers les acteurs, ou plus économiques), notre démarche étant avant tout de montrer que la dynamique du système complexe qu'est la ville, peut s'exprimer en termes de règles assez « naturelles » de transition spatiales, sans avoir recours à des systèmes d'équations complexes au paramétrage souvent obscur. De plus, différentes méthodes d'observation de la dynamique du système aident l'utilisateur à affiner son modèle puis à le valider. Ce genre d'exercice nous a semblé être un préalable important avant d'appréhender l'automate cellulaire comme un outil d'aide à la décision et de prospective. Les véritables enjeux se situent à ce niveau. En effet, les aménageurs et les politiques ont besoin d'avoir une vision à court et à moyen terme de leur territoire. L'automate cellulaire peut répondre à cette attente, non pas en proposant une seule image du territoire observé à une échéance donnée, mais en proposant différents scénarios d'évolution qui aideront le décideur dans ses choix d'aménagement.

## 5 Références bibliographiques

Amblard F. Phan D. Modélisation et simulation multi-agents pour les Sciences de l'Homme et de la Société : une introduction, Londres, Hermes-Sciences, 414 p, 2006.

Ascher, F. Ces événements nous dépassent, feignons d'en être les organisateurs. Essai sur la société contemporaine. La Tour d'Aygues, éditions de l'Aube, 304p, 2001.

Batty M., Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals, MIT Press, 648p, 2005

Benenson I, Torrens, P., Geosimulation ; Automata\_based modeling of urban phenomena, J.Wiley, 287p, 2004.

Clarke K.C., Hoppen S. and Gaydos I., «A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area», Environment and Planning B, 24, pp.247-261, 1997

Cohen J., A coefficient of agreement for nominal scales., Educ. Psychol. Meas., 1960, 20, 27-46.



- Couclelis H., «From cellular automata to urban models: new principles for model development and implementation », *Environment and Planning B*, 24, pp.165-174,1997
- Colonna A., Distephano V., Lombardo S., Papini L., Rabino A., «Learning cellular automata : modelling urban modelling », *Proceedings of the 3rd International conference on GeoComputation*, 1998.
- Delorme Marianne, *An introduction to cellular automata*, Research report n°98-37, ENS Lyon, LIP, 1998.
- Dubos-Paillard E., Guermond Y., Langlois P., « Analyse de l'évolution urbaine par automate cellulaire : le modèle SpaCelle », *L'espace géographique*, Tome 32, vol 4, pp 357-378 (22), 2003
- Guermond Y., Delahaye D., Dubos-Paillard E., Langlois P., «From modelling to experiment », *GeoJournal*, vol. 59, iss. 3, pp. 171-176 (6), Kluwer Academic Publishers, 2004
- Hagen, A., « Multi-method assessment of map similarity. » In Ruiz, M., Gould, M., Ramon, J. (Eds.) *Proceedings of the Fifth AGILE Conference on Geographic Information Science*, pp. 171–182. 2002
- Hagen-Zanker A., Straatman B., Uijee I., «Further developpments of a fuzzy set map comparaison approach », *International Journal of Geographical Information Science*, Vol 19, N° 7, pp 769-785, August 2005.
- Langlois P., « Les automates cellulaires pour la modélisation des systèmes spatiaux », in Y. Guermond (dir), *Modélisations géographie, Déterminismes et complexités*, Chap. 12, pp 321-352 (32), Ed Hermès, 2005
- Langlois P., Delahaye D. « RuiCells, automate cellulaire pour la simulation du ruissellement de surface ». *Revue internationale de géomatique*, vol 12, n° 4, p. 461-487, 2002.
- Ollinger Nicolas, *Automates cellulaires : structures*, Thèse de l'école normale supérieure de Lyon, 2002
- Pontius Jr. R.G, «Quantification error versus location error in comparison of categorical maps», *Photogrammetric Engineering and Remote Sensing* 66 (8), pp. 1011–1016, 2000.
- Weisbuch G., *Dynamique des systèmes complexes, une introduction aux réseaux d'automates*, Editions du CNRS, 1989
- White R., Engelen G. and Uljee I., «The use of Constrained Cellular Automata for high-resolution modelling of urban land use dynamics », *Environment and Planning B*, 24, pp.323-343, 1997.