

Les mobilités quotidiennes à Delhi à partir de données multi-sources : Méthodes, limites et perspectives.

Alexandre Cebeillac¹, Eric Daudé¹, Shankare Gowda²

¹ Université de Rouen, UMR IDEES

² Centre for Social Science and Humanities, Delhi

Introduction

L'analyse et la modélisation des mobilités intra-urbaines intéressent de nombreux secteurs, de la planification des transports publics, à la compréhension de phénomènes sociaux en passant par le contrôle des épidémies, notamment de maladies vectorielles comme la dengue. Cette dernière est transmise par des moustiques du genre *Aedes* (*Albopictus* et *Aegypti*), particulièrement adaptés aux villes des zones inter-tropicales.

Le virus est potentiellement mortel et concerne chaque année plusieurs millions de personnes, dont plusieurs milliers à Delhi (Inde), mégapole d'environ 26 millions d'habitants. Le virus se transmet de manière bi-directionnelle, c'est-à-dire qu'il passe (via des piqûres) d'un moustique infecté vers un Homme sain et d'un Homme contaminé vers un moustique sain. Ce mode de transmission implique que les mobilités, essentiellement humaines, ont un rôle primordial dans la propagation de l'épidémie, que cela soit à l'échelle nationale (Gatrell, 2011; Wesolowski et al., 2015), internationales (Morens and Fauci, 2013; Nunes et al., 2014; Tatem et al., 2006), ou locale et micro-locale (Stoddard et al., 2013, 2009).

Ce travail s'inscrit dans un programme de recherche plus large (Daudé, Éric et al., 2015), qui vise la mise en place d'un système multi-agent prenant en compte les dynamiques des moustiques (Maneerat, 2016), de l'environnement (Misslin, 2017) et des mouvements de populations à l'échelle de la ville de Delhi (Inde), afin de simuler et appréhender la propagation de la Dengue. Dès lors, étudier les mobilités urbaines nécessite d'avoir les concepts, méthodes, et données adaptées à l'échelle d'une ville ou d'un quartier.

Concernant les concepts, le mouvement de la time-geography, développé par Torsten Hägerstrand, a apporté à partir des années 70 la notion d'espace d'activité, « soit la portion de l'espace urbain qu'un individu fréquente lors de ces activités quotidiennes » (Horton and Reynolds, 1971). Il s'agit de l'ensemble des lieux fréquentés habituellement par un individu, et ces activités sont définies à la fois dans le temps (agenda) et dans l'espace (domaine de mobilité). Il est alors possible d'observer d'un point de vue individuel les différents types de lieux fréquentés en fonction de l'agenda d'une personne. D'un point de vue agrégé, cela permet d'étudier le potentiel d'attraction de certains secteurs selon leur fonction (activités qui s'y réalisent) en fonction du temps, et d'étudier éventuellement les différents niveaux de coprésences entre groupe sociaux.

L'avantage d'un tel concept est qu'il est très compatible avec la modélisation à base d'agent (Banos, 2013). Il est alors possible de générer des agents réalisant des activités dans différents secteurs de la ville, à différents moments, selon diverses règles et tendances de déplacement. Néanmoins, pour qu'un tel modèle soit réaliste, il convient de collecter des données de mobilités adaptées à la ville étudiée.

Delhi est une mégapole faite de discontinuités socio-économiques très marquées. Il suffit parfois de traverser une avenue pour passer de quartiers (*colonies*) informels (*slum*, *unplanned colony*, etc.) défavorisés à des quartiers planifiés où vivent des classes moyennes plus aisées.

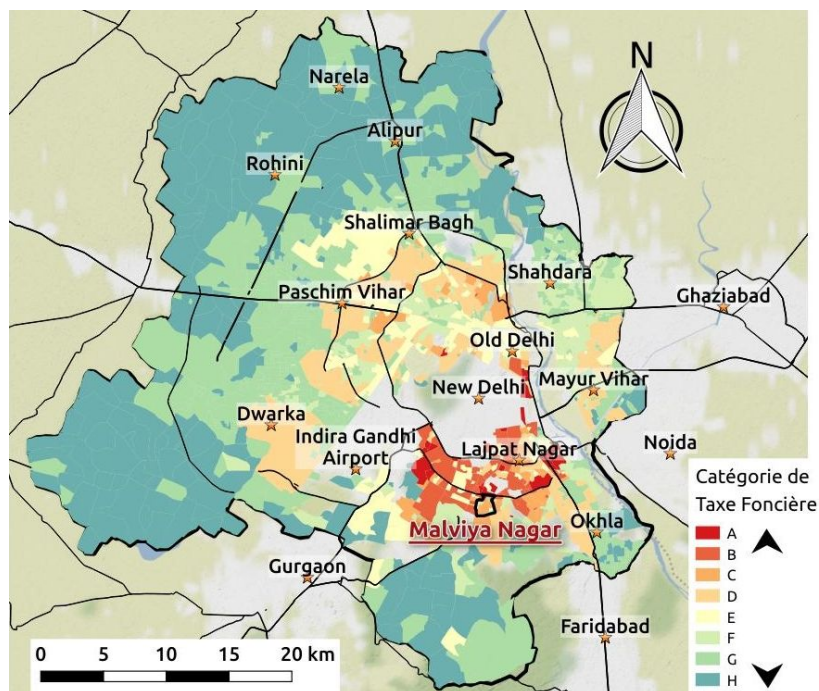


Figure 1 : Les différentes colonies à Delhi, selon leur catégorie de taxe foncière.

Le niveau de la taxe foncière dont s'acquittent les habitants des colonies dépend de l'accès aux différents services et infrastructures de bases (eau, électricité, routes goudronnées, accès aux transports, présence de parcs et d'écoles, etc.). Elle varie de la catégorie A (~quartiers les plus favorisés) à H (~villages urbains excentrés et défavorisés) et semble être un bon indicateur du niveau de ressources des habitants (Lefebvre, 2011; Telle, 2011).

Aussi, les différences dans les potentiels de déplacement entre individus (ou motilité (Kaufmann and Jemelin, 2004)) selon les secteurs de la ville, les types de quartiers (du *slum* aux *gated community*), les niveaux de ressource, le genre ou encore le capital culturel méritent d'être étudié, ne serait-ce que pour mieux comprendre les déterminants des mobilités urbaines à Delhi. Ainsi, un modèle de mobilité devrait essayer de prendre en compte ces facteurs afin d'être le plus réaliste possible.

Ce papier, à vocation exploratoire et méthodologique se décline en deux sections. La première porte sur l'acquisition des données de mobilité, qu'elles soient collectées directement sur le terrain ou qu'elles proviennent du réseau social Twitter. Dans l'optique d'une utilisation dans un modèle de mobilité alimenté par des données de différentes sources, ces données individuelles devront avoir une structure similaire permettant de décrire l'espace d'activité de chaque individu.

La seconde section évaluera dans quelles mesures le quartier de domicile peut être un indicateur du niveau de ressource économique et s'il existe des potentiels de mobilités différents selon le quartier de résidence.

1 Collecte des données de mobilité à Delhi

1.1. Enquêtes de terrain

a) Zone d'étude

Une enquête de terrain pilote a été menée dans un secteur du sud de Delhi, Malviya Nagar (figure 1). Ce dernier a la particularité d'être un assemblage de la plupart des types de quartiers que l'on trouve à Delhi et toutes les catégories de taxe foncière y sont représentées (sauf les catégories H, figure 2), dans un espace relativement restreint. La zone est très accessible, avec au nord la rocade intérieure et au sud la rocade extérieure. Le secteur est desservi par de nombreuses lignes de bus et par une ligne de métro (ligne jaunes, deux arrêts). Il est également traversé du nord au sud par une grande avenue, où se trouve un grand marché. Au sud de la zone se trouve également le *mall* de Saket. Il n'y a en donc pas d'entraves structurelles aux mobilités, et tous les types de commerces sont présents.

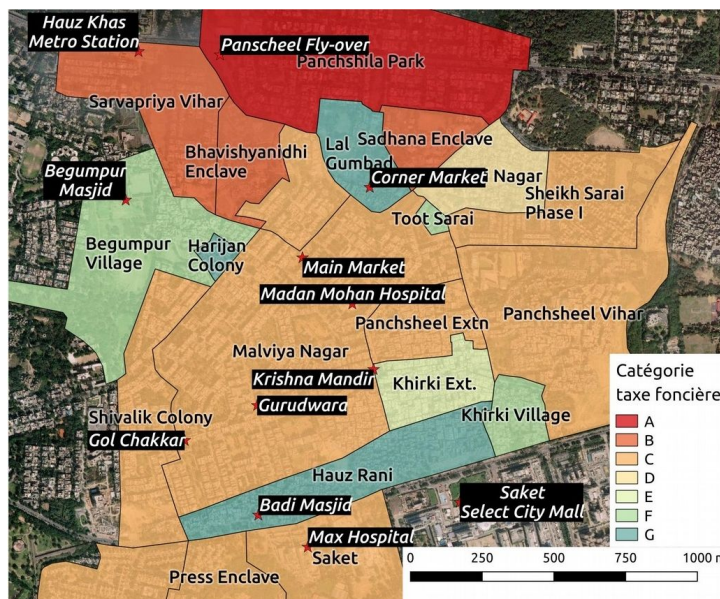


Figure 2 : Les différents quartiers du secteur de Malviya Nagar, et leurs catégories de taxe foncière.

b) Protocole de collecte

Il existe plusieurs manières de collecter des informations sur les mobilités des personnes. Il est par exemple envisageable de demander à l'interviewé de nous communiquer les différents lieux qu'il a fréquentés par tranche horaire la veille, comme c'est le cas pour les enquêtes ménage-déplacement, ou lors de la semaine précédente. Mais ces entretiens sont assez longs et peuvent exclure des lieux fréquentés de manière plus anecdotique. Nous avons ici fait le choix d'utiliser un questionnaire semi-directifs, demandant à chaque personne si elle réalise certaines activités (son travail, aller au restaurant, dans des *malls*, des marchés, des parcs, des lieux de cultes, etc.), dans quelles zones de la ville, à quel moment de la journée et à quelle fréquence hebdomadaire. Nous allons ainsi dans le sens des préconisations de (Perkins et al., 2014) sur la collecte de données de mobilité pour la simulation. À partir de ces informations, il est possible de reconstruire directement un espace d'activité individuel, caractérisé par un ensemble d'activités réalisées dans différents lieux, à différentes plages horaires et associées à des fréquences de visites.

99 entretiens semi-directifs ont été réalisés dans les différentes colonies de la zone, allant du *slum* de Begumpur à la *gated-community* de Panschilla Park. L'objectif de cette étude pilote est ici d'obtenir un échantillon composé d'un large spectre socio-économique possible, sans chercher la représentativité. Durant la durée de l'enquête (20 jours), nous n'avons pas pu interviewer autant de femmes (20) que d'hommes (79). Aucune des personnes habitants dans les zones les plus aisées (catégories A et B) n'a désiré répondre à notre questionnaire.

c) La taxe foncière, un bon indicateur de richesse économique ?

Lors des entretiens, nous demandions les revenus du foyer. Mais toutes les personnes interrogées n'étaient pas en mesure de nous fournir cette information. Nous avons donc défini un indicateur de ressource, en nous basant sur les possessions matérielles du foyer (tableaux 1).

Critère	Classe	Effectifs
climatiseur > 0 & Voiture > 1	Très élevé	9
climatiseur > 0 & 0 < Voiture <=1	Élevé	17
climatiseurs = 0 & voiture > 0 ou climatiseurs > 0 & voiture =0	Moyen	19
cooler > 0 ou motorbike > 0	Faible	29
Rien	Très Faible	18

Tableau 1 : Estimation des niveaux de ressources selon des critères de possession.

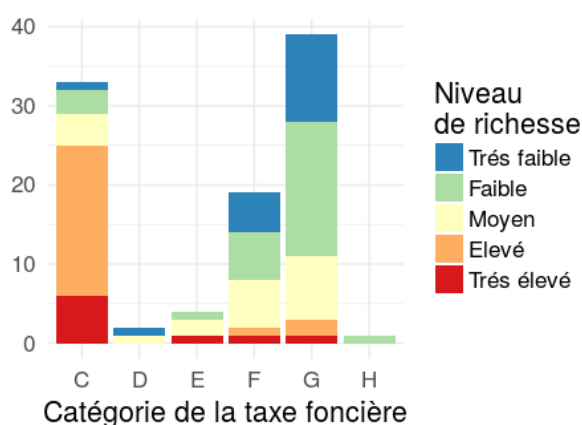


Figure 3 : Part des personnes interviewées selon la catégorie de taxe foncière de leur quartier de domicile et du niveau de ressources estimés.

Si nous mettons au regard les niveaux de richesses estimés précédemment à la catégorie de la taxe foncière du domicile (figure 3), nous pouvons noter que la plupart des personnes les plus aisées (rouge et orange) vivent dans des colonies de catégorie C. Néanmoins, ceci n'est pas systématique, car les colonies les moins favorisées (de E à G) accueillent également des personnes ayant un niveau de richesse estimé assez important. Néanmoins, cette rapide étude ne remet pas en cause la pertinence de la taxe foncière comme indicateur (partiel) de ressource économique, malgré le risque d'erreur écologique que cela implique.

1.2. Données Twitter

a) Présentation du service

En complément de ces données collectées *in situ*, nous mobilisons également des sources de données relativement récentes et accessibles : les traces numériques publiques et géolocalisées issues du réseau social Twitter.

Cette plateforme de « micro-blogging » est utilisée mensuellement par environ 300 millions de personnes à travers le monde. Elle permet aux utilisateurs du service d'envoyer des messages courts (140 caractères puis 280 à partir de 2018), par défaut visibles par tous, auxquels l'utilisateur peut inclure sa géolocalisation précise (celle de son terminal téléphonique). Ces données sont assez massives et facilement accessibles par une API¹. Elles ont déjà permis par le passé de quantifier les déplacements à l'échelle nationale et internationale (Hawelka et al., 2014; Jurdak et al., 2015) et urbaine (Cebeillac et al., 2018; Khan et al., 2017; Luo et al., 2016).

b) Collecte des données

Nous avons enregistré ici les données publiques et géolocalisées des utilisateurs de Twitter en Inde, en limitant l'extraction des données au pseudonyme, l'horodatage et les coordonnées précises générées lors de l'envoi d'un tweet. Ne sont conservés que les utilisateurs ayant envoyés des tweets dans la région de Delhi. Ce travail de collecte, de sauvegarde et de traitement de données est en cours d'enregistrement au registre de la CNIL. La période d'enregistrement des données s'étale du 25 juin 2014 au 4 décembre 2015 pour un total de 527 jours. 63 495 utilisateurs ont envoyé au moins un *tweet* dans la région de Delhi, pour un total de plus de 3,16 millions de messages.

c) Qui sont les utilisateurs de Twitter ?

Nous n'avons pas d'information sur l'âge des utilisateurs de Twitter à Delhi, mais certaines études estiment que le service concerne surtout les 15-25 au Royaume-Uni (Sloan et al., 2015), et que les hommes de moins de 30 ans sont surreprésentés à Londres (Longley et al., 2015). Notre travail se déroulant dans le contexte de la dengue, il est important de préciser que cette maladie touche surtout les jeunes et notamment les moins de 30 ans à Delhi (Cecilia, 2014), même si les plus jeunes sont les plus vulnérables. Ainsi, les utilisateurs de Twitter feraient donc partie de la tranche d'âge la plus concernée par la dengue.

d) Filtre des données

Nous avons dans un premier temps nettoyé la base de donnée en filtrant parmi les tweets envoyés ceux qui proviennent de bots (automatisation d'envois de tweets par des algorithmes) en supprimant les utilisateurs ayant envoyé plus de 2 % de leurs messages à des vitesses supérieures à 600 km/h. Nous écartons également les utilisateurs les moins réguliers, qui ont tweeté lors de moins de 15 jours sur la période. Pour plus de détails sur la méthode appliquée aux données Twitter enregistrées à Bangkok, voir par exemple Cebeillac et al., (2018, 2017).

¹ Application Program Interface. <https://dev.twitter.com/streaming/overview>

e) Construction de l'espace d'activité

Regroupement des messages dans des tweets dans des « lieux »

Nous avons créé pour chaque individu un espace d'activité, en regroupant les tweets envoyés depuis des lieux proches géographiquement. Pour cela, nous utilisons l'algorithme *Dbscan* (Ester et al., 1996), en appliquant un rayon de 50 m, comme présenté dans Cebeillac and Rault (2016). Chaque utilisateur se voit donc attribuer un ensemble de lieux, caractérisés par un nombre de tweets envoyés lors de différents jours de la semaine et à des heures données. Nous obtenons ainsi leur espace de vie « virtuel », auquel il convient d'inférer une activité potentiellement réalisée.

Définition des activités réalisées

Afin d'associer un motif de visite à chaque lieu, nous mobilisons des informations sur l'utilisation du sol en partant du principe que le type de lieu depuis lequel un individu tweet caractérise l'activité qu'il est en train de réaliser. Par exemple si une personne envoie des messages depuis un *mall*, nous considérons qu'elle est en train de visiter une galerie marchande.

N'ayant pas de données institutionnelles précises et adaptées à notre disposition, nous avons créé une couche d'utilisation du sol en utilisant d'abord des données surfaciques provenant de la plateforme de cartographie libre et participative OpenStreetMap² (OSM). Sont récupérés notamment les zones commerciales, les lieux d'éducatons (écoles, universités, bibliothèques), les hôpitaux, les lieux de cultes, les parcs, les routes ou encore les gares et aéroports.

Afin de compléter ces informations, nous utilisons également des données surfaciques de Google Maps. Le service montre en effet des cartes avec un code couleur précis³ renvoyant à des *Aera of Interest* (AOI, soit des zones d'intérêt, notamment commercial), des lieux d'éducatons, des parcs ou encore des hôpitaux. Nous avons pour cela adapté la fonction *get_map* de la librairie *ggmap* (Kahle and Wickham, 2013) disponible sous R, qui accède à l'API de Google Static Map⁴, et extrait les catégories en fonction des codes couleurs.

Détection du domicile et de l'activité principale

Parmi tous les lieux de l'espace d'activité, nous avons cherché à définir le lieu de domicile et celui où se déroule l'activité principale. Pour le lieu de domicile, nous partons du principe que l'utilisateur émet un grand nombre de messages lors de plusieurs jours, et surtout entre 20 h et 08 h. Ce lieu ne doit pas être associé à des *malls*, des parcs ou des routes. Nous effectuons le même raisonnement pour l'activité principale, sauf qu'un grand nombre de messages doivent être envoyés en journée, entre 7 h et 19 h, et qu'il ne s'agit pas du lieu de domicile.

Nous obtenons ainsi 4089 utilisateurs dans la zone élargie de la capitale Indienne (pour 758 552 tweets) dont 634 ont une activité principale. Seulement 88 personnes résident à Malviya Nagar.

2 <https://www.openstreetmap.org/>

3 <https://blog.google/products/maps/discover-action-around-you-with-updated/>

4 <https://developers.google.com/maps/documentation/maps-static/intro>

f) Représentativité de l'échantillon Twitter à Delhi

Le nombre d'utilisateurs de Twitter dont nous avons pu estimer la localisation du domicile est plutôt faible au regard de la taille de Delhi. Ces lieux de résidences sont surtout dans le sud de la ville, dans les quartiers les moins denses. L'échantillon n'est donc pas représentatif spatialement. Si nous comparons la proportion de la population de Delhi et les utilisateurs de Twitter vivant dans une colonie d'une taxe foncière données (figure 4), nous pouvons noter une surestimation des usagers de Twitter dans les zones les plus aisées (catégorie A à D) et une sous-estimation dans les colonies les plus pauvres (de F à H). Ceci suggère que les personnes qui utilisent Twitter font partie des franges les plus aisées de la population.

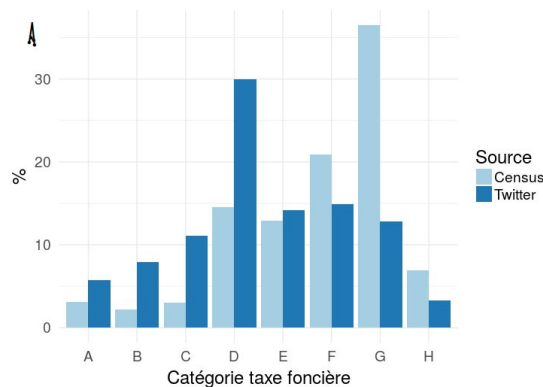


Figure 4 : Comparaison de la part de la population du recensement de 2011 et du nombre d'utilisateurs de Twitter selon la catégorie de la taxe foncière de leur lieu de domicile.

g) Rendre compatible les données Twitter et les données du terrain

Après ces différents traitements, nous obtenons pour chaque utilisateur une liste d'activités, réparties dans différents lieux, et associé à des temporalités de visite.

Il est ensuite possible d'estimer une fréquence de visite en divisant simplement le nombre de jours où une personne a tweeté dans un lieu par le nombre de jours où la personne a été active sur le réseau social. Par exemple, si une personne a envoyé des messages pendant 100 jours différents, dont 20 dans un lieu, ce dernier lieu aura une probabilité d'être visité un jour donné de 20 %. Nous pouvons aussi pondérer par le nombre de jours où une personne a tweeté depuis son lieu de domicile, en partant du principe que cette personne y est présente tous les jours. Si l'utilisateur a tweeté lors de 50 jours depuis son lieu de résidence estimé, la probabilité de visite un jour donné du lieu où elle a tweeté 20 fois peut être alors de 40 % ($100 \times 20 / 50$). Diverses méthodes sont donc envisageables, mais nous n'avons pas de moyen de validation, ni de savoir dans quelles mesures les espaces d'activités des utilisateurs de Twitter sont conformes à l'espace réellement vécu par la personne.

2 Analyse des potentiels de mobilités

Nous allons dans cette section comparer quantitativement les potentiels de mobilités des divers membres de l'échantillon. Un des objectifs est d'explorer si des tendances de déplacements sont observables en fonction de la catégorie du quartier de résidence (données Twitter) ou du niveau de ressource des membres de l'échantillon (données terrains). Pour cela, nous allons définir 3 critères de mobilités : le nombre de lieux fréquentés, le nombre d'activités effectuées et le rayon de giration – soit un indicateur de dispersion qui revient schématiquement à calculer la

moyenne des distances parcourues par rapport à un centre de gravité (González et al., 2008), défini ici comme le domicile. Nous appliquons ensuite un algorithme de partitionnement automatique, le kmeans. Nous choisissons 3 classes pour les données issues du terrain et 4 pour les données Twitter.

2.1. Données in situ

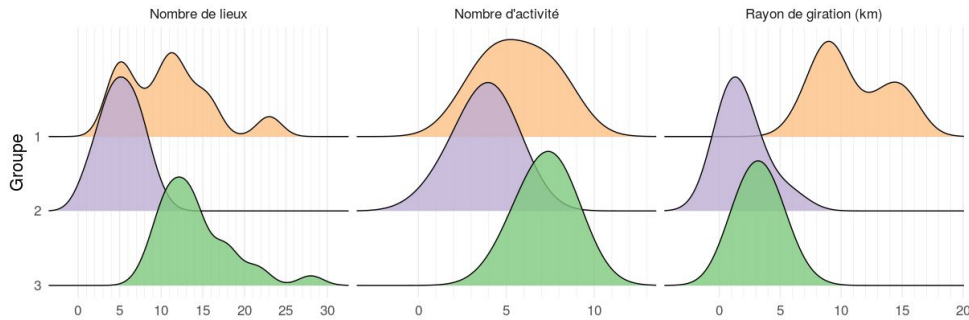


Figure 5 : Répartition des effectifs (terrain) pour les 3 classes selon les différentes métriques utilisées, présentés sous forme de fonction de densité.

La figure 5 présente pour chacune des 3 classes du kmeans les distributions des effectifs selon les trois critères de mobilités évoqués précédemment. Nous pouvons noter que le groupe 1 est caractérisé par un fort potentiel de dispersion, tandis que le nombre de lieux et d'activités ne semble pas être si discriminant (amplitude importante en x). Ce groupe est majoritairement composé d'hommes, de personnes âgées de plus de 20 ans, excluant les personnes ayant le moins de ressources économiques (figure 6). Ces dernières sont très représentées dans le groupe 2, où les femmes sont majoritaires, et qui est caractérisé par un faible nombre de lieux visités et d'activités effectuées et un rayon de giration très bas. Il s'agit des personnes les moins mobiles d'après ces indicateurs. Enfin le dernier groupe est équilibré entre homme et femme, avec une surreprésentation des personnes les plus aisées. Il s'agit des membres de l'échantillon qui ont le plus d'activités dans un grand nombre de lieux, et qui se déplacent sur des distances intermédiaires (entre le groupe 1 et 2).

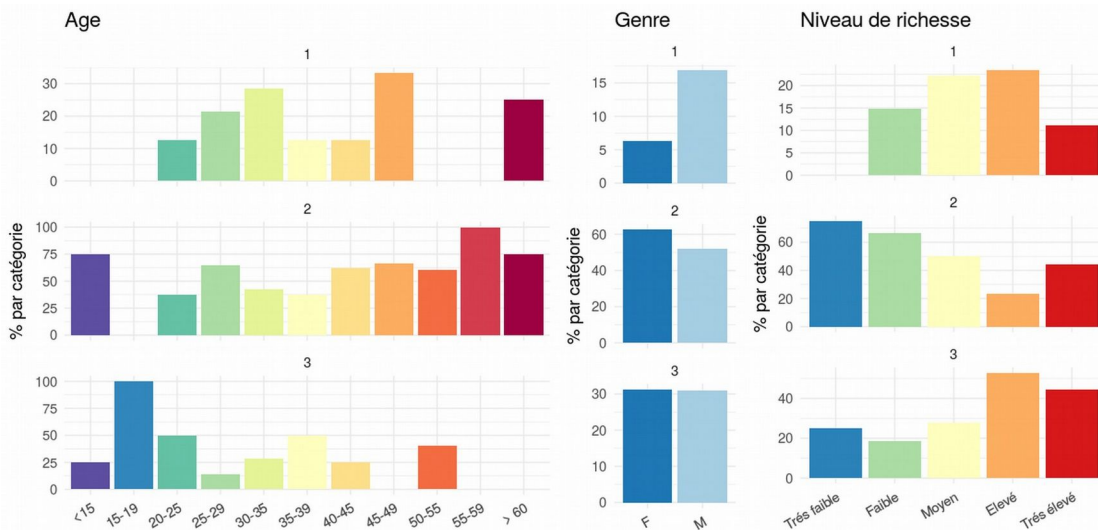


Figure 6 : Répartition des membres de l'échantillon (terrain) selon les groupes auxquels ils appartiennent, en fonction de l'âge, du genre et du niveau de ressource.

2.2. Données Twitter

Nous effectuons le même travail avec les données issues de Twitter (figure 7), mais en prenant la catégorie du lieu de domicile comme indicateur de ressources économiques (figure 8).

Les membres de la classe 1 sont ceux qui effectuent le plus d'activités et qui fréquentent le plus grand nombre de lieux. Ils ont également un rayon de giration relativement élevés par rapports aux classes 2 et 3. Ils concernent surtout les personnes habitant dans des zones assez aisées (catégories B, C et dans une moindre mesure A), mais également dans des zones de catégorie H. Les membres de la classe 2, au regard de leurs traces numériques, sont les moins mobiles, que cela soit en termes de potentiel de dispersion, de nombre de lieux fréquentés ou d'activités différentes effectuées. Ils sont répartis de manière assez homogène dans toutes les catégories de colonie. Nous pouvons noter pour les personnes appartenant à la classe 3 une légère sous-représentation des personnes résidant dans des colonies de D à H. Ils sont aussi caractérisés par un nombre d'activités assez élevés et un rayon de giration plus important que pour les membres de la classe 2. Enfin, les personnes appartenant à la classe 4 ont surtout la particularité de fréquentés des lieux assez éloignés les uns des autres (rayon de giration important). Cette dernière classe concerne surtout des personnes vivant des colonies moins favorisées, de D à H.

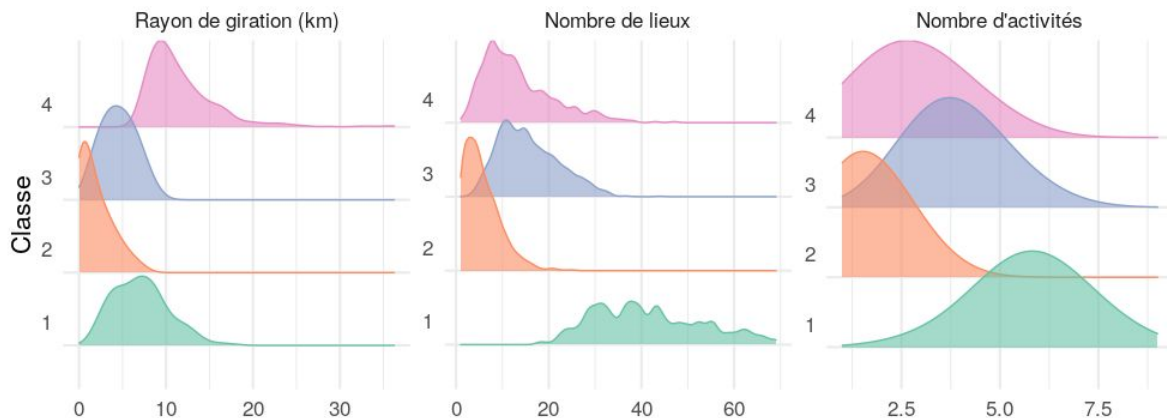


Figure 7 : Répartition des effectifs (Twitter) pour les 4 classes selon les différentes métriques utilisées, présentés sous forme de fonction de densité. Les classes 1, 2, 3 et 4 ont respectivement 323, 1271, 770, et 204 membres.

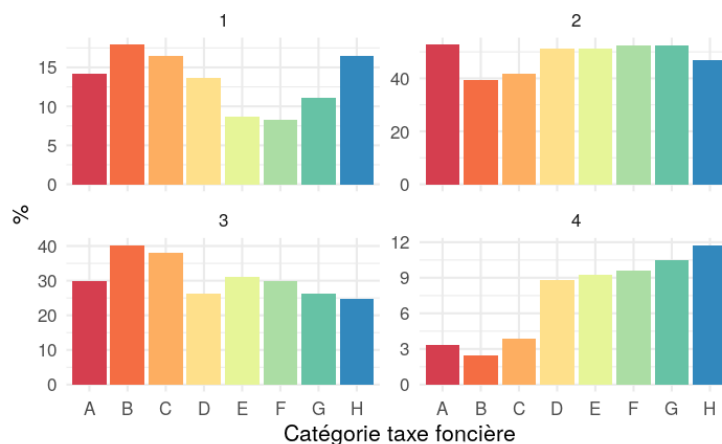


figure 8: Part des personnes résidant dans une colonie de catégorie donnée, selon le groupe résultant du kmeans.

Conclusion et perspectives d'utilisation en modélisation

L'utilisation d'un kmeans sur trois indicateurs de mobilités (rayon de giration, nombre de lieux fréquentés et nombre d'activités effectuées) permet de mettre en évidence des potentiels de mobilités différents selon les niveaux de ressources ou le genre (terrain) ou du quartier de domicile (Twitter). Il convient néanmoins de rester prudent dans l'analyse des résultats compte tenu de la faible taille des échantillons. D'autres études de terrains mériteraient d'être effectuées à Delhi, où dans des zones où les données Twitter sont plus représentatives de la population (*i.e* Bangkok (Cebeillac et al., 2018)).

Les données que nous avons collectées sur le terrain et via Twitter sont structurées de la même manière (activité, lieu, heures et fréquences de visite). Elles peuvent servir à l'élaboration d'un modèle de mobilité individu-centré à base d'agent. Ainsi, en s'inspirant des données collectées, il est possible de définir des groupes d'agents selon des potentiels de mobilités différents comme un nombre de lieux fréquentés, des types et un nombre d'activités effectuées, ainsi qu'une propension à se déplacer plus ou moins importante. Ces groupes pourraient notamment dépendre de la zone géographique ou d'un capital socio-économique estimé à partir de la localisation de leur domicile.

Cela dit, les membres de l'échantillon n'ont pour l'instant qu'un espace d'activité brut, non ordonné dans le temps, et il conviendra par la suite de leur recréer un agenda continu dans le temps, soit succession d'activités réalisées à différents moments de la journée. Pour cela, nous emploieront une approche probabiliste, en attribuant des jours et des heures de visites en nous basant sur les fréquences de visites. Il devient alors possible d'extraire des informations à partir des agendas reconstitués pour générer des agendas de synthèses. Par exemple, l'utilisation de chaîne markovienne permet de poser une probabilité de réaliser une activité à une heure donnée sachant l'activité que l'agent est en train de réaliser à l'heure précédente (Wu et al., 2014). Ces méthodes employées et les divers résultats obtenus sont détaillées dans Cebeillac, (2018) et feront l'objet de future(s) publication(s).

Néanmoins, nous sommes conscients que le volume de données Twitter à Delhi est extrêmement faible et non représentatif spatialement. Ce n'est pas le cas à Bangkok, une autre zone d'étude de notre projet, où cette méthodologie peut également être mise en place. Aussi, notre approche dépend plus de la structure des données que de leur source. Il est donc aussi envisageable d'utiliser des données issues par exemples des statistiques d'appels et pour ensuite appliquer des traitements similaires.

Bibliographie

- Banos, A., 2013. Pour des pratiques de modélisation et de simulation libérées en Géographie et SHS. Paris.
- Cebeillac, A., 2018. Mobilités urbaines et données en ligne, pour l'étude des maladies vectorielles à Delhi (Inde) et Bangkok (Thaïlande). Université de Rouen, Rouen.
- Cebeillac, A., Daudé, Éric, Vaguet, Alain, 2017. Discontinuités spatiales, santé et mobilités. Analyses et typologies de Google POI et de Tweets pour caractériser les structures spatiales et les dynamiques d'attractivités de Bangkok (Thaïlande). Presented at the Sageo, Rouen.

- Cebeillac, A., Huraux, T., Daudé, É., 2018. Where? When? and how often? What can we learn about daily urban mobilities from Twitter data and google map in Bangkok (Thailand) and what are the perspectives for dengues studies? *Netcom Réseaux Commun. Territ.*
- Cebeillac, A., Rault, Y.-M., 2016. Contribution of geotagged Twitter data in the study of a social group's activity space. The case of the upper middle class in Delhi, India. *Netcom Réseaux Commun. Territ.* 231–248.
- Cecilia, D., 2014. Current status of dengue and chikungunya in India. *WHO South-East Asia J. Public Health* 3, 22. <https://doi.org/10.4103/2224-3151.206879>
- Daudé, Éric, Vaguet, Alain, Paul, Richard, 2015. La dengue, maladie complexe. *Nat Sci Soc* 23, 331–342. <https://doi.org/10.1051/nss/2015058>
- Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *KDD-96 Proc.*
- Gatrell, A.C., 2011. *Mobilities and Health*. Ashgate (Surrey).
- González, M.C., Hidalgo, C.A., Barabási, A.-L., 2008. Understanding individual human mobility patterns. *Nature* 453, 779–782. <https://doi.org/10.1038/nature06958>
- Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., Carlo, R., 2014. Geo-located Twitter as the proxy for global mobility patterns. *Cartogr. Geogr. Inf. Sci.*
- Horton, F.E., Reynolds, D.R., 1971. Effects of Urban Spatial Structure on Individual Behavior. *Econ. Geogr.* 47, 36. <https://doi.org/10.2307/143224>
- Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., Newth, D., 2015. Understanding human mobility from Twitter. *PloS One* 10, e0131469.
- Kahle, D., Wickham, H., 2013. ggmap: Spatial Visualization with ggplot2. *R J.* 5.
- Kaufmann, V., Jemelin, C., 2004. La motilité, une forme de capital permettant d'éviter les irréversibilités socio-spatiales ? Presented at the *Espaces et sociétés aujourd'hui. La géographie sociale dans les sciences et dans l'action*, Rennes.
- Khan, S.F., Bergmann, N., Jurdak, R., Kusy, B., Cameron, M., 2017. *Mobility in Cities: Comparative Analysis of Mobility Models Using Geo-tagged Tweets in Australia*.
- Lefebvre, B., 2011. *Les services hospitaliers de Delhi : planification, privatisation et gouvernance urbaine*. Université de Rouen, Rouen.
- Longley, P.A., Adnan, M., Lansley, G., 2015. The geotemporal demographics of Twitter usage. *Environ. Plan. A* 47, 465–484.
- Luo, F., Cao, G., Mulligan, K., Li, X., 2016. Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago. *Appl. Geogr.* 70, 11–25. <https://doi.org/10.1016/j.apgeog.2016.03.001>
- Maneerat, S., 2016. *Modélisation à base d'agents des risques vectoriels en milieux urbains: exemple d'Aedes aegypti, vecteur de la dengue, à Delhi (Inde)*. Université de Rouen, Rouen.
- Misslin, R., 2017. *Modélisation de l'environnement d'un moustique vecteur de maladies - L'exemple d'Aedes aegypti à Delhi (Inde) et Bangkok (Thaïlande)*. Université de Rouen, Rouen.
- Morens, D.M., Fauci, A.S., 2013. Emerging Infectious Diseases: Threats to Human Health and Global Stability. *PloS Pathog.* <https://doi.org/10.1371/journal.ppat.1003467>
- Nunes, M.R.T., Palacios, G., Faria, N.R., Sousa, E.C., Jr, Pantoja, J.A., Rodrigues, S.G., Carvalho, V.L., Medeiros, D.B.A., Savji, N., Baele, G., Suchard, M.A., Lemey, P., Vasconcelos, P.F.C., Lipkin, W.I., 2014. Air Travel Is Associated with Intracontinental Spread of Dengue Virus Serotypes 1–3 in Brazil. *PLoS Negl Trop Dis* 8, 1–13. <https://doi.org/10.1371/journal.pntd.0002769>
- Perkins, T.A., Garcia, A.J., Paz-Soldan, V.A., Stoddard, S.T., Reiner, R.C., Vazquez-Prokopec, G., Bisanzio, D., Morrison, A.C., Halsey, E.S., Kochel, T.J., Smith, D.L., Kitron, U., Scott, T.W., Tatem, A.J., 2014. Theory and data for simulating fine-scale human movement in an urban environment. *J. R. Soc. Interface* 11, 20140642–20140642. <https://doi.org/10.1098/rsif.2014.0642>

- Sloan, L., Morgan, J., Burnap, P., Williams, Ma., 2015. Who Tweets? Deriving the demographic characteristics of age, occupation and socialClass from Twitter User Meta-Data. *PloS ONE*. <https://doi.org/doi:10.1371/journal.pone.0115545>
- Stoddard, S.T., Forshey, B.M., Morrison, A.C., Paz-Soldan, V.A., Vazquez-Prokopec, G.M., Astete, H., Reiner, R.C., Vilcarrromero, S., Elder, J.P., Halsey, E.S., Kochel, T.J., Kitron, U., Scott, T.W., 2013. House-to-house human movement drives dengue virus transmission. *Proc. Natl. Acad. Sci.* 110, 994–999. <https://doi.org/10.1073/pnas.1213349110>
- Stoddard, S.T., Morrison, A.C., Vazquez-Prokopec, G.M., Paz Soldan, V., Kochel, T.J., Kitron, U., Elder, J.P., Scott, T.W., 2009. The Role of Human Movement in the Transmission of Vector-Borne Pathogens. *PLoS Negl. Trop. Dis.* 3, e481. <https://doi.org/10.1371/journal.pntd.0000481>
- Tatem, A.J., Rogers, D.J., Hay, S.I., 2006. Global Transport Networks and Infectious Disease Spread, in: *Advances in Parasitology*. Elsevier, pp. 293–343.
- Telle, O., 2011. *Aedes: Analyse de l'émergence de la dengue et simulation spatiale*. Université de Rouen.
- Wesolowski, A., Qureshi, T., Boni, M.F., Sundsøy, P.R., Johansson, M.A., Rasheed, S.B., Engø-Monsen, K., Buckee, C.O., 2015. Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc. Natl. Acad. Sci.* 112, 11887–11892. <https://doi.org/10.1073/pnas.1504964112>
- Wu, L., Zhi, Y., Sui, Z., Liu, Y., 2014. Intra-Urban Human Mobility and Activity Transition: Evidence from Social Media Check-In Data. *PLoS ONE* 9, e97010. <https://doi.org/10.1371/journal.pone.0097010>