

Apports et limites des données OpenStreetMap pour l'analyse des centralités commerciales dans les espaces transfrontaliers

Marianne Guérois, Ronan Ysebaert, Timothée Giraud, Béatrice Maranget, Grégory Hamez, Mathias Boquet, Nicolas Dorkel

▶ **To cite this version:**

Marianne Guérois, Ronan Ysebaert, Timothée Giraud, Béatrice Maranget, Grégory Hamez, et al.. Apports et limites des données OpenStreetMap pour l'analyse des centralités commerciales dans les espaces transfrontaliers. [Rapport de recherche] RIATE. 2020, 57 p. hal-03587399

HAL Id: hal-03587399
https://hal.science/hal-03587399

Submitted on 24 Feb 2022

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.



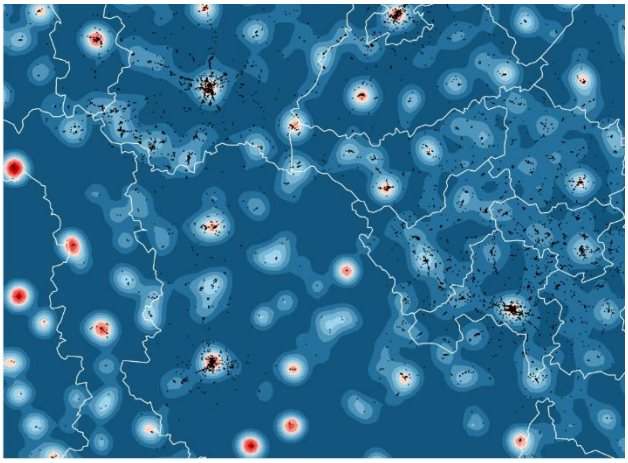
Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

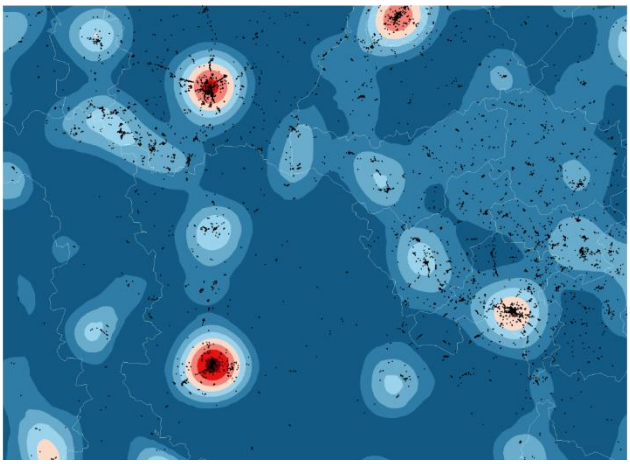
Apports des données *OpenStreetMap* pour l'analyse des centralités commerciales dans les espaces transfrontaliers

Étude ANCT – UMS RIATE – Octobre 2020

Marianne Guérois, Ronan Ysebaert, Timothée Giraud, Béatrice Maranget (Université de Paris, CNRS, RIATE)

Grégory Hamez, Mathias Bocquet, Nicolas Dorkel (Université de Lorraine, LOTERR)





Suivi de projet assuré par Gilles Toutin (ANCT)

Table des matières

Introduction 4

1. Quelle information géographique sur les équipements commerciaux dans la base *OpenStreetMap* ? 8

1.1. La qualité des données POI d'OSM en question 8

1.2. L'information géographique sur les équipements commerciaux dans OpenStreetMap 10

1.3. Quelle correspondance entre les tags d'OSM et les catégories d'équipement de la BPE ? 13

1.4. Extraction et mise en forme des données OSM 17

2. Analyse de complétude d'OSM à l'échelle régionale (partie française) : comparaison avec la base BPE de l'INSEE 19

2.1. Mesurer les écarts OSM/BPE d'après les quantités communales d'équipements 19

2.2. Mesurer les écarts OSM/BPE d'après la diversité des équipements présents dans les communes 25

3. Analyse de complétude d'OSM à l'échelle locale : comparaison avec des relevés de terrain dans 6 zones tests (France, Allemagne) 34

3.1. Choix des sites d'enquête, collecte et mise en forme des relevés de terrain 34

3.2. Résultats : écarts globaux et par catégorie d'équipement 40

4. Identifier des pôles de concentration d'équipements OSM à l'échelle transfrontalière 43

4.1. Mesurer des potentiels d'équipements dans le voisinage des lieux 43

4.2. Analyse de la relation entre les potentiels locaux d'équipements et de population 46

4.3. Identifier des pôles de concentration relative d'équipements 48

Conclusion 54

Bibliographie 55

Table des figures

Figure 1 - Caractéristiques démographiques de l'espace d'étude 6
Figure 2 - Les 15 équipements retenus pour l'étude (nomenclature BPE) 6
Figure 3 - La complétude des POI OSM comparés aux POI de Navteq, à l'échelle des wards du Grand Londres (Mashadi et al., 2015) 10
Figure 4 – Gestion des objets géographiques et attributs dans la base OpenStreetMap 11
Figure 5 – Exemple de « tags » proposés dans le Wiki OSM pour la clé = « amenity » 12
Figure 6 – Description du tag « amenity=restaurant » et proposition de clés-valeurs OSM qui peuvent être associées à cet équipement 15
Figure 7 – Correspondances de nomenclatures BPE-OSM-CNFG 16
Figure 8 – Extraction depuis OSM des équipements d'intérêt 17
Figure 9 – Nombre d'équipements et écarts OSM-BPE, par catégorie d'équipement 21
Figure 10 – Répartition communale des équipements et variation spatiale des écarts OSM-BPE (exemple des équipements de la gamme de proximité) 22
Figure 11 – Taux de complétude d'OSM dans les communes françaises de l'espace d'étude, par rapport à la BPE (pour 15 équipements) 22
Figure 12 – Nombre d'équipements et écarts OSM-BPE, par catégorie de commune 23
Figure 13 – Nombre d'équipements et écarts OSM-BPE, par classe de taille des communes 24
Figure 14 – Présence comparée d'équipements dans OSM et la BPE (cafés-restaurants et banques) 25
Figure 15 – Nombre d'équipements différents présents dans les communes de l'espace d'étude 27
Figure 16 – Seuils d'apparition et de généralisation de l'équipement alimentation 28
Figure 17 – Seuils d'apparition pour les équipements, comparaison OSM-BPE 29
Figure 18 – Localisation des communes en fonction de leur profil d'équipement 31
Figure 19 – Les 6 zones d'enquête sélectionnées 35
Figure 20 – Relevés de terrain dans trois villes françaises et deux villes allemandes 36
Figure 21 – Relevés de terrain, création d'une zone tampon et extraction des 15 catégories d'équipements dans OSM et dans la BPE : exemple dans l'hypercentre de Metz 37
Figure 22 –Problèmes liés à la méthode de la zone tampon 38
Figure 23 –Problèmes de définition de l'appartenance aux catégories 39
Figure 24 –Problèmes de décalages temporels OSM/BPE ou de mise à jour de la BPE 39
Figure 25 – Dénombrement des équipements recensés par terrain d'enquête, source de données et équipement 41
Figure 26 – Densité des équipements OSM pour 1000 habitants, à l'échelon des NUTS3 (pour les cafés-restaurants, banques, pharmacies, magasins d'électronique) 44
Figure 27 – Densité d'équipements à l'échelon local (pour les cafés-restaurants, banques, pharmacies, magasins d'électronique) 44
Figure 28 – Potentiel d'équipements OSM dans un voisinage de 5 km (cafés-restaurants, banques, pharmacies, magasins d'électronique 45
Figure 29 – Modèle de régression linéaire entre le nombre d'équipements et la population résidente dans un voisinage de 5km. 47
Figure 30 – Résidus des modèles de régression linéaire entre le potentiel d'équipements et de population dans un voisinage de 5km. 48

Figure 31 – Pics relatifs de centralité : rapport entre les potentiels d'équipements mesurés entre 2 portées (pour les cafés-restaurants, banques, pharmacies, magasins d'électronique) 49
Figure 32 – Pics locaux de concentration d'équipements – 10 % 50
Figure 33 – Pics locaux de concentration d'équipements – 5 % 51
Figure 30 – Pics régionaux de concentration d'équipements – 10 % 51
Figure 35 – Pics régionaux et locaux de concentration d'équipements, tentative de croisement des approches 52

Tableau 1 - Nombre d'équipements présents d'après OSM et la BPE (15 équipements de l'étude) ... 26
Tableau 2 - Diversité des équipements présents dans les communes de l'espace d'étude 27
Tableau 3 - Reconstitution de gammes d'équipements à partir des seuils d'apparition dans OSM/BPE 29
Tableau 4 - Niveaux d'équipements présents dans les communes françaises de la zone d'étude, pour une sélection de 7 équipements 30
Tableau 5 - Le nombre d'équipements (15 catégories) dans les 6 zones tests, selon les sources 40

Introduction

La répartition spatiale des équipements commerciaux et de services, un enjeu de dialogue transfrontalier

L'organisation territoriale des commerces en France métropolitaine fait l'objet d'un intérêt renouvelé depuis une dizaine d'années, du fait d'une restructuration profonde des commerces de proximité ayant affecté en particulier les villes petites et moyennes ainsi que les bourgs et villages des espaces ruraux (Barcszak et Hilal, 2016 ; Baudet-Michel, 2018). Tout en s'inscrivant dans ces dynamiques générales, les espaces transfrontaliers qui sont au cœur de cette étude sont affectés par des évolutions spécifiques : pour les habitants, l'ouverture des frontières a encouragé les pratiques transfrontalières et multiplié les opportunités d'accès aux équipements des territoires voisins, en exploitant les avantages comparatifs de contextes nationaux différenciés (Smits, 2006 ; Hamez, de Ruffray, 2009). Pour les commerçants, ces dynamiques peuvent alimenter des situations de concurrence et de saturation commerciale. Parallèlement, des projets de coopération institutionnelle se construisent pour faciliter l'accès aux ressources commerciales et de services au-delà des frontières, dans un contexte où domine encore la prédominance de cultures nationales d'aménagement commercial (Devillet et al., 2012). L'identification des équipements et des centralités de commerces apparaît pour certains comme un enjeu important de dialogue transfrontalier.

Des données sur les équipements très lacunaires à l'échelle transfrontalière

Toutefois, il existe peu de travaux sur la répartition des équipements commerciaux à l'échelle transfrontalière : l'observation statistique et cartographique de ces phénomènes se heurte en effet à l'obstacle majeur de la disponibilité et de l'hétérogénéité de données comparables. En France, leur analyse repose très largement sur la Base Permanente des Equipements (BPE) de l'INSEE, qui est la base de données institutionnelle de référence pour les travaux portant sur les commerces et services. Mais il s'agit d'une exception à l'échelle européenne : au-delà des frontières, il reste très difficile d'obtenir des données équivalentes pour l'ensemble des équipements, à l'échelle nationale ou même régionale. L'analyse de l'équipement des territoires en commerces et services reste un angle mort de la comparaison des territoires européens, même si ces lacunes sont partiellement comblées par la réalisation d'enquêtes sur des portions de territoires (Devillet et al., 2012) ou par des études portant sur un type d'équipement (par exemple les maternités, De Ruffray, Hamez, 2009), au prix d'un long travail d'harmonisation des bases de données.

OSM, une base alternative pour la comparaison internationale des équipements commerciaux ?

Dans ce contexte, la base de données contributive *OpenStreetMap* présente une ressource potentiellement riche d'informations géographiques sur les équipements commerciaux. Plusieurs travaux ont déjà discuté et souligné l'intérêt du contenu de cette base pour analyser les infrastructures de réseaux routiers. OSM contient aussi des données sur de nombreux autres éléments visibles de l'espace public, notamment sur les équipements commerciaux et de services qui font partie des points d'intérêt (POI, en anglais *points of interest*) décrits dans la base. Mais jusqu'à présent, ces données ont été peu exploitées dans une perspective analytique, à l'exception de quelques travaux, tels que par exemple le projet *Prophecy*

d'ESPON (ESPON 2017[1]), qui s'est appuyé sur OSM pour étudier l'implantation territoriale de 9 services d'intérêt général dans les 32 pays couverts par le programme ESPON (médecins, hôpitaux, pharmacies, banques, écoles primaires, écoles secondaires, supermarchés, alimentation, cinémas). Plusieurs travaux ont pourtant souligné la richesse de l'information collectée et sa complémentarité avec les données institutionnelles, notamment sur les réseaux routiers. Qu'en est-il des données sur les équipements commerciaux ?

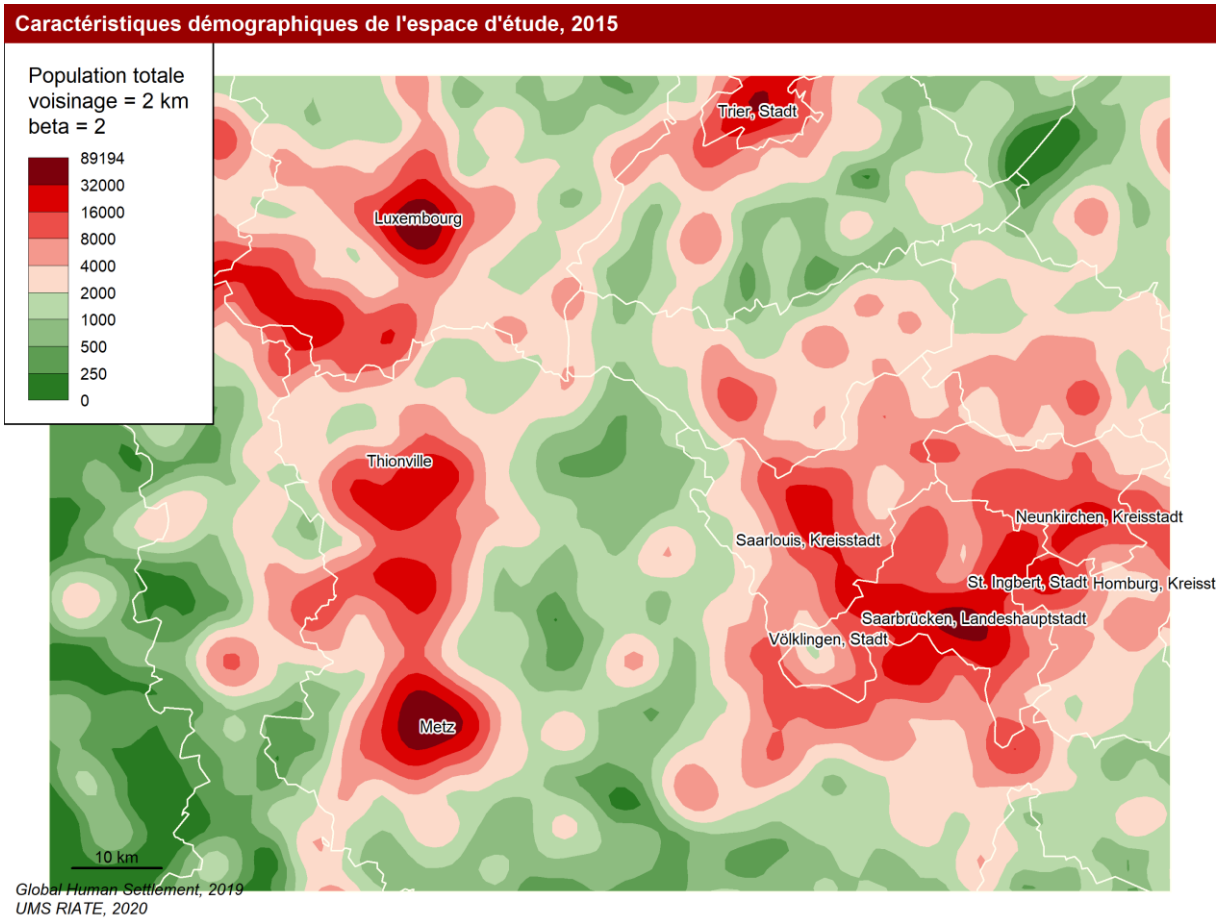
L'objectif de cette étude de nature méthodologique est de documenter l'apport potentiel d'une base de données contributive telle qu'*OpenStreetMap* (OSM) pour analyser l'organisation territoriale des équipements commerciaux au sein d'espaces transfrontaliers. Elle présente à ce titre un double intérêt : contribuer aux travaux encore peu nombreux sur les équipements commerciaux à l'échelle transfrontalière et surtout, contribuer à l'analyse encore rare de la complétude d'OSM sur les données d'équipements. D'un point de vue méthodologique, elle repose sur une analyse de la complétude des données OSM menées par comparaison avec différentes données de référence, à différentes échelles : d'une part la BPE (INSEE) à l'échelle régionale côté français, d'autre part un ensemble de relevés de terrain effectués par l'équipe du LOTERR (Université de Lorraine) dans plusieurs sites localisés en France et en Allemagne. Cette analyse de complétude doit permettre d'évaluer la faisabilité d'une comparaison de la répartition des équipements commerciaux à l'échelle transfrontalière, sur la base d'une sélection des équipements OSM les mieux contribués.

Une étude de cas sur la Grande Région, un protocole d'analyse reproductible

L'espace d'étude défini pour ce projet est un espace transfrontalier de dimension régionale, couvrant le nord de la région Grand-Est en France, le sud du Luxembourg, l'ouest de la Sarre et de la Rhénanie-Palatinat en Allemagne (*Figure 1*). Cette vaste zone, qui représente environ 12800 km² et 2.97 millions d'habitants (en 2011, d'après Eurostat), comprend sur ses marges quatre grandes aires urbaines fonctionnelles (FUA) dont l'emprise recouvre partiellement l'espace d'étude : Luxembourg au nord[2], Metz au sud (370 000 hab. en 2016, d'après la base des FUA d'Eurostat), Sarrebrück à l'Est (800 000 hab.), Trèves au nord-est (260 000 hab.). Elle offre une grande variété de situations d'encadrement territorial par les commerces (Devillet et al., 2012). Si le choix d'un espace d'étude particulier se justifie par l'intégration de relevés de terrain, l'ensemble de la méthodologie mise en œuvre pour l'étude, depuis l'extraction des données OSM jusqu'à l'identification de pôles de concentration d'équipements commerciaux, est conçue de manière à pouvoir être transposée à d'autres espaces transfrontaliers situés entre la France et les Etats voisins. Les différentes étapes du protocole sont ainsi documentées dans des fiches techniques accessibles en ligne (https://frama.link/centralite), qui reprennent l'ensemble des programmes d'analyse statistique et cartographique construits et implémentés dans le logiciel R, en les commentant de manière détaillée.

[1] https://riate.cnrs.fr/wp-content/uploads/2017/07/InnerPeripheries_09072018.pdf.
[2] Les données de FUA pour Luxembourg sont indisponibles.

Figure 1 - Caractéristiques démographiques de l'espace d'étude



Une sélection de 15 équipements

L'analyse s'appuie sur la sélection d'un jeu d'équipements considérés dans la littérature comme pertinents pour l'analyse des centralités commerciales, en tant que marqueurs de différents niveaux de rayonnement, depuis les commerces de proximité jusqu'aux équipements plus rares correspondant à la gamme supérieure de la BPE (*Figure 2*).

Figure 2 - Les 15 équipements retenus pour l'étude (nomenclature BPE)

Equipements	Gamme BPE
Alimentation Boucherie Boulangerie Coiffeur Pharmacie Restaurants/Cafés	Proximité
Ameublement Banque Chaussures Electronique Opticien Stations-service Vêtements	Intermédiaire
Cinéma Théâtre	Supérieur

La constitution de cette liste répond à différents critères de sélection qui s'appuient d'une part sur la littérature existante, d'autre part sur la disponibilité des données dans OSM. Ainsi, 9 de ces équipements ont été choisis sur la base de travaux scientifiques récents : les commerces d'alimentation, les boucheries, les boulangeries, les pharmacies, les magasins d'ameublement, d'électronique, de chaussures, de vêtements, figurent notamment parmi les équipements sélectionnés par les auteurs de l'étude sur la rétraction des commerces et services dans les villes petites et moyennes (Baudet-Michel et al., 2019). Un équipement (opticiens) a été ajouté du fait de son caractère discriminant dans l'identification d'un niveau

de centralité intermédiaire par Hilal et al. (2020), dans le cadre de l'étude « Petites centralités » mandatée par l'ANCT. 5 autres équipements (banques, cafés-restaurants, cinémas, stations-service, théâtres) ont été retenus du fait de l'importance des contributions les concernant dans OSM. A l'inverse, nous avons écarté des équipements pertinents *a priori* mais dont la qualité de description dans OSM était insuffisante au regard des besoins de l'analyse : c'est le cas notamment des moyennes et grandes surfaces commerciales, dont les différents niveaux hiérarchiques (supérettes, supermarchés, hypermarchés…) ne pouvaient être distingués dans les données OSM[3]. De plus, certains des 15 équipements retenus présentent un intérêt particulier en tant qu'équipements « transfrontaliers » (restaurants et cafés, stations-services…), susceptibles d'être fréquentés par les personnes habitant de des deux côtés de la frontière.

Déroulé du rapport

Ce rapport s'articule autour de quatre parties :

- Nous présentons tout d'abord la structure et la nature de l'information sur les équipements commerciaux dans d'*OpenStreetMap* (OSM) (partie 1). Au-delà de la question essentielle du degré d'exhaustivité de cette base de données contributive, nous nous interrogerons sur la manière dont les catégories d'équipements sont décrites et sur les possibilités de correspondance avec les nomenclatures hiérarchisées des bases institutionnelles ou professionnelles de référence.
- Premier volet de l'analyse de complétude, la partie 2 repose sur la comparaison systématique des données OSM avec les données de référence de la BPE de l'INSEE (2018) dans les 703 communes françaises de la zone d'étude et pour la sélection des 15 équipements d'intérêt. A travers la mesure des écarts OSM-BPE, appréhendés aussi bien en termes de présence/absence que de quantité d'équipements recensés dans les communes, il s'agit d'évaluer le degré de complétude de la base OSM et sa variation selon les catégories d'équipements et les types d'espaces, des centres-villes aux communes rurales isolées.
- Dans la partie 4, ce sont des relevés de terrain réalisés dans plusieurs sites en France et en Allemagne par l'équipe du LOTERR qui tiennent lieu de données de référence et au regard desquels la complétude des données OSM (et BPE en France) est évaluée, à l'échelle infra-communale. Ces enquêtes permettent d'élargir l'analyse de la qualité d'OSM à la partie allemande de la zone d'étude, et constituent des points de repère précis en l'absence de base de référence institutionnelle.
- Enfin, partant des équipements pour lesquels les données OSM s'avèrent les plus fiables d'après les différents volets de l'analyse de complétude, la partie 5 s'attachera à analyser la répartition spatiale des équipements OSM dans l'ensemble de l'espace transfrontalier et à identifier des pôles de centralité définis par une concentration relativement élevée d'équipements.

[3] La sélection d'équipements de service public a été évoquée au début du projet, mais la description de l'information correspondante (sur les écoles et les équipements de santé notamment) ne permettait pas de distinguer les niveaux hiérarchiques essentiels (par exemple la distinction entre écoles primaires, collèges et lycées) dans la qualification des centralités.

1. Quelle information géographique sur les équipements commerciaux dans la base *OpenStreetMap* ?

La base *OpenStreetMap* (OSM) compte parmi les sources pionnières et les plus connues d'Information Géographique Volontaire (VGI). Elle offre à ce titre une "alternative intéressante à l'information géographique générée par des professionnels, lorsque cette dernière est trop coûteuse, qu'elle n'existe pas pour une thématique donnée, ou bien que des mises à jour temporelles sont nécessaires[4]" (Goodchild, 2007). Base contributive et libre de droits, issue d'un projet créé en 2004, OSM a vocation à représenter l'ensemble des "phénomènes visibles à la surface de la terre" et à ce titre est souvent présentée comme le "Wikipedia de la cartographie". Parmi la large palette de données collectées (https://www.openstreetmap.fr/), les points d'intérêt (*Points of Interest*, POI) désignent des lieux remarquables pour les usagers, tels que des lieux touristiques, des équipements (écoles, stations de transport…), des magasins. Si ces POI constituent avant tout une ressource majeure pour un grand nombre d'applications en ligne, que ce soit pour construire des itinéraires, enregistrer sa localisation (*check-in*), ou encore poster des photos géolocalisées, ces données sont aussi de plus en plus exploitées dans une perspective de recherche (Jonietz, Zipf, 2016). En témoigne leur exploitation pour analyser la dispersion de la population à des échelles fines (Bakillah et al., 2014), identifier des lieux d'activités et de sociabilité en ville (Bawa-Cavia, 2011 ; Vienne et al., 2017) ou encore, à des échelles plus vastes, pour comparer l'équipement des régions européennes en services élémentaires à la population (projet Prophecy d'ESPON, 2017).

En raison de la nature contributive de la base, d'un mode de production de données décentralisé et non supervisé, la plupart des travaux scientifiques consacrés à OSM accordent une très large place à la question de la qualité de ces données. Il existe ainsi une vaste littérature sur la qualité des données de réseaux routiers ou d'occupation du sol, un peu moins volumineuse et moins connue sur les POI. Dans cette partie, nous reprendrons les principaux résultats de ces travaux avant de présenter plus précisément la nature de l'information sur les équipements commerciaux présente dans OSM. Nous exposerons ensuite les difficultés soulevées par la comparaison des POI d'OSM avec les autres bases utilisées dans cette étude (la base institutionnelle de la BPE, les relevés de terrain), du fait de l'hétérogénéité des nomenclatures, et la nécessité de construire une table de correspondance entre catégories d'équipement. Enfin, nous expliquerons comment ont été mises en œuvre l'extraction et la mise en forme des données OSM pour sélectionner les 15 équipements d'intérêt à l'échelle de la zone d'étude.

1.1. La qualité des données POI d'OSM en question

De nombreuses caractéristiques des données OSM découlent de leur caractère « librement éditable et librement utilisable (…), aucun système de contrôle des informations ne séparant l'édition par un contributeur de l'utilisation par quelqu'un d'autre » (Hayat, 2019). En l'occurrence, la question de la qualité de la base, a été abordée dans de nombreuses publications scientifiques et synthétisée dans plusieurs travaux récents (voir par exemple Hombiat et al., 2016 ; Hayat, 2019). Les **premiers travaux sur les données de réseaux**

[4] "An attractive alternative for use cases where professional geographic information is too expensive, not available for a theme, or when timely updates are required" (Goodchild, 2007)

routiers se sont focalisés sur l'évaluation de la précision de la localisation des objets, de la complétude et de la justesse attributaire des données (Viry et al., 2015). Ils ont montré qu'au-delà d'un certain seuil de contributions, les jeux de données issus d'OSM étaient en général comparables aux données géographiques traditionnelles maintenues par les agences nationales et que ce degré de complétude augmentait avec le temps (Zielstra et Zipf, 2010). Dans certains cas, les données de VGI sont jugées suffisamment « matures » pour pouvoir être considérées comme des données de substitution ou d'enrichissement des données conventionnelles. Ces constats sont cependant relativisés par la forte hétérogénéité spatiale de la complétude des données OSM, à différentes échelles : à l'échelle régionale, cette complétude est élevée dans les agglomérations, plus faible dans les espaces ruraux (Hacklay, 2010, Zielstra ; Zipf, 2010), l'information se dégradant aussi bien du point de vue du nombre de POI que de l'exactitude des attributs qui les décrivent (nom des routes, par exemple). A l'échelle internationale, d'autres résultats sur les réseaux routiers ont mis l'accent sur la variation de l'activité des communautés OSM selon les pays (Viry et al., 2015).

Par comparaison, les **travaux consacrés à la qualité des POI** sont moins nombreux et portent avant tout sur la validation des méthodes d'évaluation de la qualité, tandis que les expérimentations sur des jeux de données restent plus limitées. Touya et al. (2017) ont ainsi proposé une synthèse des méthodes d'évaluation de la qualité des POI, rappelant la distinction classique entre les **méthodes basées sur les caractéristiques internes de la base** (historique des contributions et rectifications, relations topologiques entre les objets OSM) et les **méthodes basées sur la comparaison avec des données « externes »**, qui peuvent être soit des données conventionnelles, soit d'autres données de VGI. Jonietz et al. (2016) se sont focalisés sur les méthodes mises en œuvre pour comparer les données contributives à des données de référence. Ils distinguent trois grands types d'indicateurs : les **indices de complétude** s'appuient sur la comparaison du nombre d'objets représentés dans chaque base ; les **indicateurs de précision géométrique** reposent soit sur des mesures de distance euclidienne entre deux points décrivant le même phénomène, soit sur l'identification des données de VGI présentes à une certaine distance (buffer) des points de référence ; enfin, la **précision thématique** peut être examinée à l'aide de différentes mesures (pourcentage de classification correcte d'objets, distance de Levenshtein[5] mesurant le degré de similarité entre deux chaînes de caractère, appliquée en général au nom des établissements). Touya et al. (2017) insistent sur le fait qu'aucune de ces méthodes ne se suffit à elle-même ; ils proposent des méthodes automatisées reposant sur la combinaison de ces différentes approches, qu'ils testent sur données de localisation des stations de métro OSM à Paris, par comparaison avec les données de la BD topo de l'IGN, en utilisant un algorithme de mise en correspondance terme-à-terme des bases[6].

Parmi les travaux les plus poussés de comparaison avec les POI d'une base de référence, Mashadi et al. (2015) ont analysé les variations de la **complétude des POI OSM dans le Grand Londres** en les comparant aux POI de Navteq. Une correspondance terme à terme entre les objets des deux bases est effectuée à l'aide de la combinaison d'un critère de proximité géographique -moins de 100 m- et d'un critère lexicographique –via la distance de Levensthein-. Ayant calculé à l'échelon des *Wards*[7] un indice de complétude compris entre 0

[5] Mesure de la différence entre deux chaînes de caractères.

[6] Pour chaque objet de la BD topo, les « candidats » à la correspondance sont sélectionnés à l'intérieur d'une zone-tampon de 350 m autour des points de référence ; puis sont évalués à partir 3 critères de ressemblance qui s'appuient sur la position –distance euclidienne- et sur le nom associé à l'objet (distance normalisée de Levenhstein sur les noms des stations de métro)

[7] Unité spatiale élémentaire de la géographie administrative britannique, issue du découpage des *Districts*.

et 1 et basé sur le rapport entre le nombre de POI OSM et le nombre de POI *Navteq*, ils concluent à une **complétude globale de l'ordre d'un tiers (0,35) pour le Grand Londres**, mais très variable selon les wards (Figure 3) : cette variation se calque en partie sur un gradient centre-périphérie, mais montre aussi une forte hétérogénéité au sein des périphéries. La plupart des wards montre ainsi une complétude proche de 0, tandis que pour certains d'entre eux, elle atteint 0,6. La probabilité de contribution est d'autant plus élevée que la population est élevée, comme le montre la corrélation positive et significative (coefficient de corrélation de 0,3) avec la densité de population, et que la localité est proche d'un des 10 centres d'emploi principaux du Grand Londres. En revanche, l'hypothèse d'une relation inverse entre complétude et pauvreté, qui supposerait que les contributions sont moindres dans les localités les plus pauvres, du fait d'un moindre accès à internet et de temps de loisir plus courts, n'est pas validée.

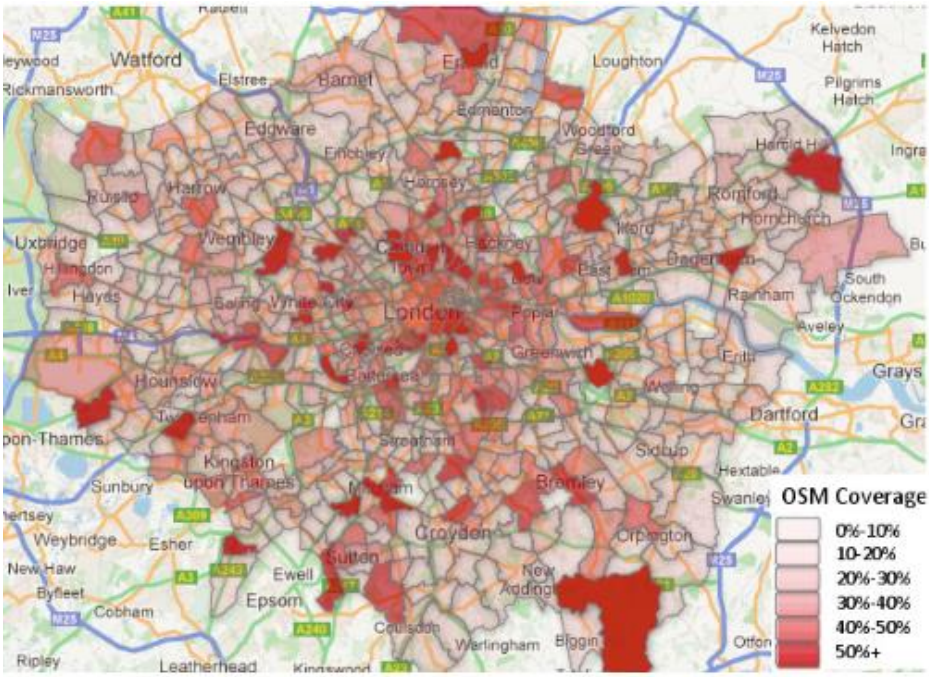


Figure 3 - La complétude des POI OSM comparés aux POI de Navteq, à l'échelle des wards du Grand Londres (Mashadi et al., 2015)

Par comparaison avec les données de réseaux routiers, les travaux portant sur la qualité des POI OSM restent peu nombreux et soulignent l'intérêt d'un approfondissement de la comparaison entre les POI OSM et ceux des bases de référence. Cela nécessite au préalable de bien comprendre comment sont collectées et structurées les données sur les équipements dans OSM.

1.2. La gestion de l'information géographique sur les équipements commerciaux dans OpenStreetMap

La manière dont la base OSM est alimentée et structurée diffère très nettement des bases de données conventionnelles sur les équipements. Dans les données contributives, les informations qualitatives sont peu structurées et sont produites par une communauté (Hayat, 2019). De fait, la base OSM est alimentée par des contributeurs de profils variés, correspondant en général à des individus particuliers, mais aussi à des associations ou à des collectivités territoriales. Les sources de saisie sont également de différents types : il s'agit la plupart du temps de relevés terrains (individuels ou organisés dans le cadre de campagnes locales), réalisés à l'échelon de petites zones, mais les contributions peuvent être aussi réalisées à partir de la digitalisation d'entités sur des images aériennes ou encore à partir de

bases de données externes et importées. Ces sources de saisie peuvent être renseignées dans les métadonnées. Dans les faits, elles ne le sont que de manière très aléatoire (par exemple, pour les boulangeries présentes dans les communes françaises de l'espace d'étude, 95% des contributions ne sont pas identifiées), sauf lorsqu'il s'agit de contributions professionnelles (c'est le cas par exemple des pharmacies).

Géométrie des objets OSM

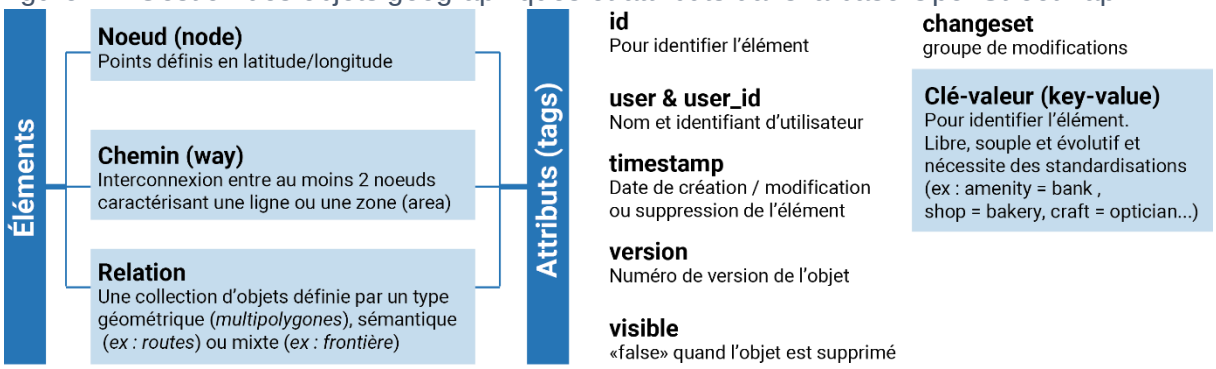
Lorsqu'un objet est ajouté à la base OSM, il est d'abord décrit à partir de ses caractéristiques géométriques élémentaires, de forme et de localisation. Chaque objet a ainsi pour élément soit un nœud (pour un objet ponctuel), soit un chemin (ouvert pour un objet linéaire, fermé pour un objet surfacique). Il peut aussi être décrit par une relation géométrique (par exemple une relation d'inclusion d'un élément ponctuel au sein d'une surface).

L'information est collectée à un niveau de granularité très fin, proche de l'expérience terrain, qui garantit une certaine précision dans la localisation. **Chaque équipement est en général représenté à l'aide d'un point localisé au niveau de l'accès principal du bâtiment**. Toutefois, il peut y avoir des **cas plus complexes**, lorsque plusieurs entrées de bâtiments ou bien une entrée et le périmètre d'un bâtiment sont saisis pour un même équipement. Dans l'étude, ces doublons ont fait l'objet d'un pré-traitement, tous les polygones associés à des équipements ont été transformés en points (centroïde du polygone), et en cas de points multiples, nous avons fait la moyenne des coordonnées pour résumer la localisation à l'aide de coordonnées uniques.

Attributs des objets OSM : une classification libre, guidée par un Wiki de bonnes pratiques

Chaque objet est caractérisé par un ensemble de « tags » (littéralement, des « étiquettes »). Pour un contributeur, les premiers tags à saisir sont des métadonnées qui renseignent sur le contexte de production de l'information (identifiant de l'élément, identité du contributeur, horodatage –pour une première saisie ou pour une modification-), que nous n'exploitons pas dans l'étude. Puis chaque contributeur peut décrire la qualité de l'objet à l'aide de tags élémentaires ou plus détaillés (tags secondaires). Chaque tag est conçu comme une paire d'attributs composée d'une « clé » (catégorie générale) et de sa « valeur » (sous-catégorie). Il existe ainsi 23 catégories générales (équipements, bâtiments, autoroutes…) auxquelles sont associées différentes sous-catégories. Par exemple à la clé « amenity » (catégorie « équipement »), pourra correspondre la valeur « shop » (« magasin »).

Figure 4 – Gestion des objets géographiques et attributs dans la base OpenStreetMap



Éléments et attributs utilisés pour l'étude : Extraction de l'ensemble des éléments répondant à 37 couples clé-valeur associés aux 15 équipements cibles. Puis transformation des éléments non ponctuels en points (centroides des polygones ou multi-polygones)

Il est important de **présenter plus précisément cette classification structurée par clés-valeurs,** dans la perspective d'une mise en correspondance avec la classification des bases institutionnelles. Dans OSM, il n'existe pas de nomenclature d'équipement à proprement parler, au sens où un équipement serait associé à une catégorie détaillée relevant de catégories plus larges, selon une hiérarchie emboîtée. Les attributs décrits par le couple clé-valeur constituent ce qui est communément appelé la « folksonomie » d'OSM, terme désignant littéralement une activité de classification (taxonomie) par les usagers (folk) et reflétant « le résultat de l'activité d'association libre d'étiquettes (tags) à des données ou objets dans le but personnel de les retrouver ultérieurement (Wal, 2005, cité par Hombiat et al., 2016) ». Volontairement, OSM ne restreint pas les « clés » (ni leurs « valeurs » assignées) des attributs ajoutés aux objets saisis et le contributeur est libre de choisir quel tag sera le plus approprié. Il n'existe pas de relation hiérarchique *a priori* entre les tags, comme pour les nomenclatures emboîtées. De plus, un même objet peut être associé à plusieurs tags : il peut y avoir plusieurs clés/valeurs pour un même élément, reflétant la diversité des fonctions associées à un équipement (par exemple une station-service qui ferait aussi office d'alimentation). Comme le soulignent Hombiat et al. (2016), cela permet une catégorisation rapide, simple et flexible des objets décrits, qui est adaptée au caractère évolutif de la base. Ces différents éléments rendent *OpenStreetMap* extrêmement flexible, mais ne facilitent pas toujours le partage et la réutilisation des informations contribuées.

La **diversité des tags n'est « pas pour autant synonyme de production chaotique d'attributs »** (Hayat, 2019). Il existe en effet une liste de référence qui recense les clés et les valeurs communément utilisées, rassemblées dans un Wiki qui en fournit une description, et « chaque contribution commence souvent par un passage en revue du wiki d'OSM » (Ibid). La figure extraite du WikiOSM (https://wiki.openstreetmap.org/wiki/Key:amenity) en montre un exemple pour la clé « amenity », avec la liste des valeurs associées (Figure 5). Le contributeur peut choisir d'utiliser les tags couramment employés ou d'en créer de nouveaux (http://taginfo.openstreetmap.org/keys), la liste n'étant **ni fermée, ni contraignante**.

Figure 5 – Exemple de « tags » proposés dans le Wiki OSM pour la clé = « amenity »

Source : https://wiki.openstreetmap.org/wiki/FR:%C3%89l%C3%A9ments_cartographiques#Consommation

Équipements (amenity)

Utilisé pour cartographier les équipements pour les visiteurs et les résidents comme: toilettes, téléphones, banques, pharmacies, écoles. Voir Amenities pour une introduction.

Clé	Valeur	Élément	Commentaire	Représentation	Photo
			Consommation		
amenity	bar		Bar, bistrot. voir aussi amenity=cafe		
amenity	bbq		Barbecue. Un gril public pour la cuisson de viandes ou de légumes, le plus souvent à l'extérieur.		
amenity	biergarten		Brasserie en plein air. Spécifique à l'Allemagne. Voir Biergarten.		
amenity	cafe		Café.		
amenity	drinking_water		Source d'eau potable pour promeneurs et randonneurs. Voir aussi natural=spring pour une source naturelle amenity=water_point pour ravitaillement en eau potable pour plaisanciers et campingcaristes amenity=fountain pour une fontaine (potable ou non).		
amenity	fast_food		Restauration rapide. Le type de nourriture servie peut être étiqueté avec cuisine=*.		

Au final, F. Hayat (2019) estime à la suite de Hall et al. (2018) que les « bonnes pratiques écrites dans le wiki ou simplement généralisées dans la pratique s'approchent du procédé classique de définition de spécification pour une classification ». La **contribution est très majoritairement cohérente avec les instructions du wiki**, même si les contributeurs négligent les tags secondaires documentés dans le wiki (horaires d'ouverture, fermeture définitive…) et si les tags d'OSM ne sont pas forcément représentatifs des données collectées : parfois le tag existe mais il n'est pas utilisé (par exemple le tag brasseries est peu utilisé et les brasseries sont souvent codées en 'cafés').

1.3. Quelle correspondance entre les tags d'OSM et les catégories d'équipement de la BPE ?

L'analyse sur l'équivalence des classifications d'équipement entre OSM et la Base Permanente des Equipements (BPE) de l'INSEE, qui tiendra lieu de principale base de référence dans cette étude, est un préalable indispensable à la comparaison des données.

La BPE, une base institutionnelle nationale, soumise à une enquête annuelle de qualité

La BPE est la **base de référence de la statistique publique française sur les équipements commerciaux et de services** (182 catégories d'équipement répertoriées). Produite depuis 2007 dans le prolongement de l'Inventaire communal, elle fait l'objet d'une mise à jour annuelle. Dans la version du 1er janvier 2018 retenue dans cette étude, la quasi-totalité des équipements (et notamment les 15 sélectionnés) est géolocalisée à l'aide de coordonnées géographiques (X,Y) avec une bonne précision (moins de 100 m d'erreur). Une grande partie des données (environ 60%) sont importées depuis la base SIRENE (Système Informatique pour le Répertoire des ENtreprises et leurs Établissements) gérée par l'INSEE. Ainsi, 12 des 15 équipements retenus dans l'étude relèvent de cette source, tandis que les données relatives aux trois autres équipements sont alimentées à partir de fichiers administratifs spécifiques (le site data.gouv.fr du ministère de l'Économie et des finances pour les stations-services, le CNC pour les cinémas, ARTCENA pour les théâtres). La description des équipements repose sur une nomenclature emboîtée qui se décline en 3 niveaux de détail : le premier niveau contient 7 domaines (dont les commerces, les services aux particuliers, l'enseignement…), le deuxième 27 sous-domaines et le troisième 188 types d'équipements. Ces équipements se répartissent en 3 gammes (de proximité, intermédiaire, supérieure) qui reflètent la plus ou moins grande rareté de leur présence sur le territoire. Enfin, si la BPE peut être considérée comme une base de référence pour évaluer la complétude d'OSM dans les communes françaises, il importe d'avoir en tête l'existence de **plusieurs biais pouvant conduire à des excédents ou déficits d'équipements**, qui sont identifiés lors des enquêtes annuelles de qualité (**encadré 1**).

Encadré 1 : Déficits et excédents d'équipements dans la Base Permanente des Equipements : enquêtes annuelles de l'Insee [8]

Chaque année, l'Insee met en œuvre une enquête de qualité de la BPE sur un échantillon d'équipements, dans les communes de 400 à 10 000 habitants concernées par la campagne du recensement de la population. L'enquête vise à mesurer d'éventuels excédents ou déficits d'équipements dans la BPE par rapport aux observations de terrain. Elle a permis d'identifier cinq facteurs de sur ou sous-représentation d'équipements :

- Le problème des **« faux actifs »** : plusieurs enregistrements excédentaires peuvent être liés au maintien de personnes ou d'entreprises qui ne sont plus en activité mais qui restent inscrites dans la base SIRENE. En effet, l'inscription dans cette base est obligatoire alors que la désinscription liée à une cessation d'activité ne l'est pas. Ce biais est particulièrement sensible pour les commerces, qui présentent un taux d'excédents supérieurs à la moyenne, notamment les plus volatils d'entre eux (boulangeries, boucheries, …). Ce phénomène peut être accentué dans certains types d'espaces (notamment dans les zones touristiques).
- Les **doublons** : les mêmes équipements peuvent parfois être enregistrés plusieurs fois : c'est le cas notamment des banques (excédent d'environ 14%), du fait des phénomènes de restructuration et de l'existence d'activités administratives internes. C'est le cas également des pharmacies qui peuvent être enregistrées plusieurs fois lorsqu'elles changent de catégories juridiques.
- La **définition des équipements** : les déficits d'équipements concernent parfois des catégories dont la définition est relativement restrictive dans la BPE. Il s'agit par exemple des théâtres, car seules les scènes nationales sont répertoriées. C'est aussi le cas des stations-services (déficit d'environ 10%), qui sont enregistrées à partir d'un seuil élevé (vente d'au moins 500 000 litres de carburant au cours de l'année n-1), qui conduit à la sous-estimation des stations dans les petites communes.
- Le **mode d'enregistrement de l'activité économique :** dans le cas d'activités multiples, seule l'activité principale est enregistrée, suivant le plus grand chiffre d'affaires déclaré. Les Hôtels-restaurants, par exemple, sont souvent classés dans la catégorie hôtels, sauf s'ils sont déclarés comme deux entités différentes.
- La difficulté d'enregistrement d'activités au **statut économique particulier** : en l'occurrence, la BPE saisit mal le phénomène des auto-entrepreneurs et des établissements déclarés qui n'ont pas pignon sur rue (par exemple les nombreux coiffeurs à domicile, ou encore certaines agences immobilières).

Une table de correspondance entre tags OSM et catégories d'équipement BPE

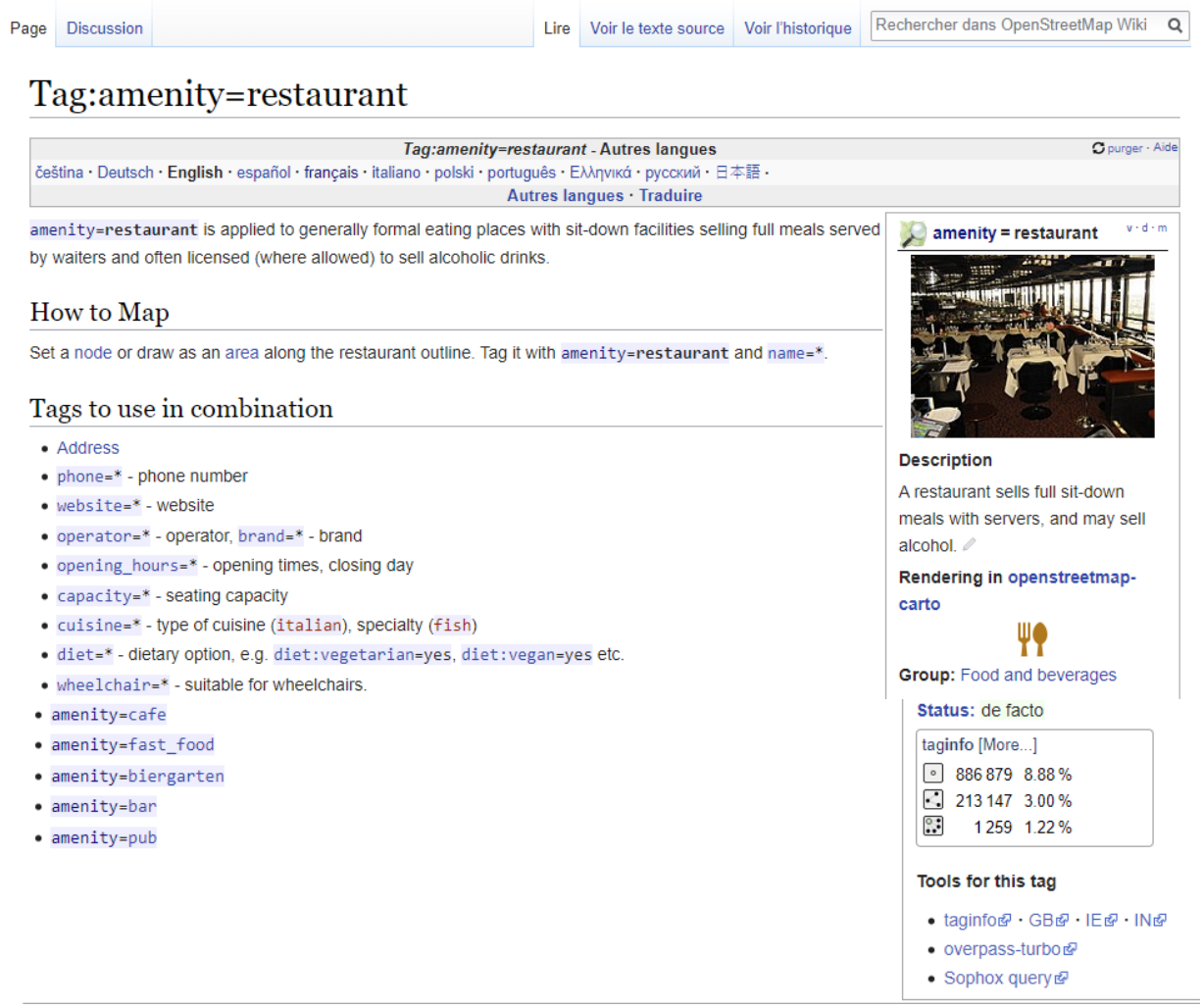
Etant données les logiques différentes de description des équipements dans OSM et dans la BPE (*folksonomie* versus nomenclature hiérarchisée), il est nécessaire d'identifier les correspondances possibles entre la nomenclature BPE et les couples clés/valeurs d'OSM afin de reconstituer une cohérence sémantique entre les deux bases. Partant des 15 catégories sélectionnées dans la BPE, une des difficultés principales tient au fait qu'il peut y avoir plusieurs couples de clé-valeurs d'OSM pour chaque catégorie de la BPE, et inversement. Afin d'identifier la ou les clés-valeurs correspondant à chaque catégorie de la BPE, la navigation sur le Wiki d'OSM (https://wiki.openstreetmap.org) a permis de rechercher les clés-valeurs correspondant à un nom d'équipement et de suggérer des tags similaires dans la nomenclature (rubrique « see also » dans la *Figure 6* ci-dessous).

[8] Les informations de cet encadré sont issues d'une part de la documentation mise en ligne sur le site de l'Insee (https://www.insee.fr/fr/statistiques/3568656#avertissements-methodologiques, également diffusée par Progedo : http://www.progedo-adisp.fr/documents/lil-1331/BPE18_Presentation.pdf), d'autre part d'échanges avec le pôle BPE de l'INSEE.

Trois cas de figure peuvent être distingués :

- Pour certains équipements, la **correspondance est directe et simple** : c'est le cas par exemple des boulangeries, des banques ou des cinémas. La clé-valeur OSM est attribuée par des contributeurs OSM sans ambiguïté et se réfère à un équipement bien identifié.
- Dans d'autres cas, la **correspondance sémantique est plus ambiguë**, au sens où une catégorie de la BPE peut être associée à plusieurs clés-valeurs dans OSM (cas par exemple des magasins d'électroménager, ou de meubles, qui sont souvent identifiés à partir de leurs marques « Brand= » dans OSM), ou à l'inverse, une même clé-valeur dans OSM correspond à plusieurs catégories dans la BPE (c'est le cas par exemple des restaurants et des cafés). Une certaine ambiguïté peut aussi être liée à la variation de la description de l'équipement selon le pays de contribution, comme en atteste par exemple la description du tag « amenity=bar » dans le wiki OSM, revenant sur la définition des bars dans le nord et le sud de l'Europe[9].

Figure 6 – Description du tag « amenity=restaurant » et proposition de clés-valeurs OSM qui peuvent être associées à cet équipement



[9] *« Un établissement commercial qui vend des boissons alcoolisées à consommer sur place (…). Généralement on n'y vend pas de véritable repas (…). Bien que les pubs (amenity=pub) soient très similaires, les bars ont généralement une ambiance plus commerciale (…). Dans les pays méditerranéens, le mot "bar" a un sens différent (ce qui ne signifie pas nécessairement que le tag doive être appliqué différemment). Ici le bar est partie intégrante du mode de vie. Vous pouvez y aller le matin pour petit-déjeuner, à midi ils servent des repas simples, toute la journée (quand ce n'est pas fermé après midi) ».*

- Parfois, enfin, il n'y **a pas de correspondance suffisante pour conserver la catégorie BPE**. C'est le cas par exemple des grandes surfaces alimentaires (la BPE distingue les supermarchés des superettes et des épiceries, ce qui n'est pas possible dans OSM), ou encore des écoles. Dans OSM, les tags associés pour les écoles sont de 4 types : school, kindergarten, university ou college. Dans la pratique, la plupart des établissements sont taggés comme des school ou university. Afin d'instaurer une hiérarchie dans les établissements, il est proposé aux contributeurs d'associer à l'établissement son ISCED d'appartenance (nomenclature européenne de référence). La plupart des contributeurs n'ayant pas connaissance de cet ISCED, cet attribut n'est de fait que très peu utilisé.

Au final, une table de correspondance a pu être établie pour 15 équipements de la BPE (fusion cafés-restaurants) (*Figure 7*). Ces catégories seront agrégées pour permettre la comparaison des données enregistrées. Dans OSM, elles regroupent 37 couples de clés-valeurs, correspondant surtout à des clés « shop » (26) et secondairement « amenity » (10), beaucoup plus rarement « craft » (1).

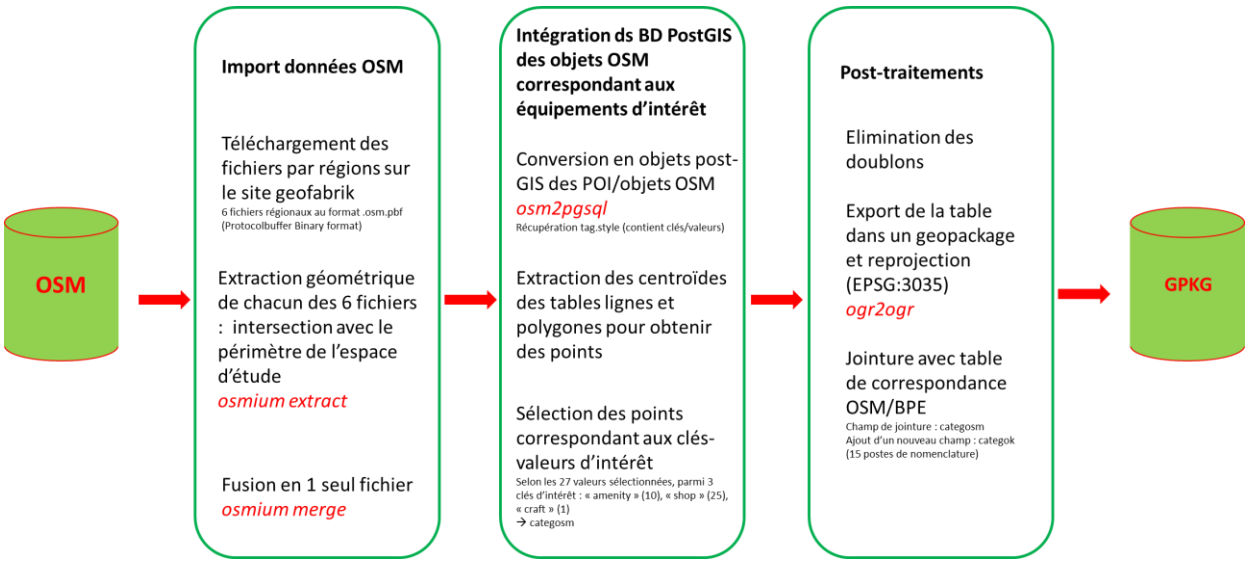
Figure 7 – Correspondances de nomenclatures BPE-OSM-CNFG

Équipement	Gamme	BPE (2018)	Tag OSM (clé/valeur)	Classement catégoriel établissement commerciaux (Commission Géographie Commerce)
Alimentation	Proximité	B201/ B202	shop=convenience /shop=greengrocer / shop=alcohol / shop =farm / shop=dairy / shop=cheese / shop=frozen_food	Alimentation générale (niveau 2) / alimentation biologique / alimentation diététique (niveau 3)
Ameublement	Intermédiaire	B306	shop=furniture /shop=antiques / shop=bathroom_furnishing / shop = lamps / shop=kitchen / shop=bed	Ameublement (niveau 2)
Banque	Intermédiaire	A203	amenity=bank	Banque (niveau 3)
Boulangerie	Proximité	B203	shop=bakery	Boulangerie – pâtisserie (niveau 3)
Boucherie	Proximité	B204	shop=butcher	Boucherie-charcuterie / volailles-gibiers (niveau 3)
Chaussures	Intermédiaire	B304	shop=shoes	Chaussures (niveau 2)
Cinéma	Supérieur	F303	amenity=cinema	Cinéma (niveau 3)
Coiffeur	Proximité	A501	shop = hairdresser	Salon de coiffure (niveau 3)
Electronique	Intermédiaire	B305	shop=electronics / shop=appliance / shop=computer / shop=hifi / shop=mobile_phone	Appareils électro-ménagers / matériel informatique / informatique / informatique (réparation / pièces) niveau 3
Opticien	Intermédiaire	B313	shop=optician / craft = optician	Optique (niveau 3)
Pharmacie	Proximité	A501	amenity=pharmacy / shop=chemist	Pharmacie (niveau 3)
Restaurant **Café**	Proximité	A504	Pour restaurant : amenity=restaurant / amenity=fast_food Pour café-bar : amenity=cafe, amenity=pub, amenity=biergarten	Restaurant rapide (fast food) / restaurant / restauration rapide / café-restaurant / hôtel-restaurant / brasserie / crêperie / friterie (niveau 3)
Station service	Intermédiaire	B316	amenity=fuel	Stations service (niveau 3)
Théâtre	Supérieur	F306	amenity=theatre	Théâtre (niveau 3)
Vêtements	Intermédiaire	B302	shop=clothes / shop=fashion	Vêtements (niveau 2)

1.4. Extraction et mise en forme des données OSM

Une fois identifiées les clés-valeurs d'OSM correspondant aux équipements d'intérêt, une méthode d'extraction et de mise en forme des données *OSM* a été mise en oeuvre. Cette démarche est synthétisée dans la *Figure 8* et la procédure est détaillée dans la ressource méthodologique associée à l'étude (préparation des données)[10].Elle s'appuie sur plusieurs outils d'extraction et de manipulation des objets de la base OSM, qui permettent une mise en œuvre relativement rapide de la chaîne de traitements, facilitant le traitement des données au sein de formats compatibles avec le langage et logiciel R.

Figure 8 – Extraction depuis OSM des équipements d'intérêt



L'importation des données a été effectuée depuis le site Geofabrik (www.geofabrik.de), qui facilite l'accès aux données OSM en les diffusant en accès libre par l'intermédiaire de fichiers découpés par sous-zones géographiques et administratives (par continent, par pays et par région), au sein de différents formats pré-traités (OSM PBG, OSM BZ2, SHP). Ces extractions sont mises à jour de façon hebdomadaire, mais il est possible également de récupérer des extractions plus anciennes du contenu d'OSM (jusqu'à 2014 pour les régions françaises). L'importation a été effectuée sur les fichiers de septembre 2019, pour les 6 régions couvrant la zone d'étude. **Cela signifie concrètement que la présentation des résultats de l'étude porte sur cette version d'OpenStreetMap,** qui a nécessairement évolué avec les contributions réalisées depuis. Dans un second temps, tous les objets ont été convertis en format PostGIS, sélectionnés par requête attributaire à partir des 27 clés-valeurs décrivant les équipements d'intérêt. Lorsque ces objets étaient d'implantation linéaire ou surfacique, leur centroïde a été récupéré de manière à obtenir une base composée uniquement d'éléments ponctuels. Tous les polygones ont été transformés en points (centroides des polygones). Leur localisation est définie par un système de coordonnées géographiques exprimé en latitude/longitude. Enfin, plusieurs opérations de post-traitement ont permis d'éliminer les doublons, de reprojeter la base (projection européenne de référence, ETRS89-LAEA) et d'effectuer une jointure attributaire avec la table de correspondance afin d'associer aux POI d'OSM les catégories de la BPE.

[10] https://rcarto.gitpages.huma-num.fr/centralite/A_preparation.html

L’ensemble de ces opérations sont documentées et peuvent être aisément transposées à tout autre espace d’étude couvert à la fois par la BPE et par OSM sur les équipements commerciaux ciblés par l’étude.

Conclusion partie 1

S’il existe une littérature importante sur la qualité des données de réseaux de transport dans OSM, cette question a été jusqu’à présent peu traitée pour les données de POI et peu testée empiriquement. Les quelques études de cas réalisées à l’échelle de quartiers, plus rarement d’agglomérations, insistent sur la précision élevée de la localisation des POI d’OSM, tout en mentionnant la forte hétérogénéité spatiale de leur complétude. Mais il n’existe pas, à notre connaissance, d’étude systématique des POI portant sur de plus vastes échelles.

Dans la perspective d’une évaluation de la complétude des POI OSM à une vaste échelle, par comparaison avec une base de données externe institutionnelle (BPE), nous avons créé une table de correspondance entre catégories d’équipements, en nous appuyant notamment sur la consultation du Wiki d’OSM. Bien que les logiques de construction des deux bases OSM et BPE soient très différentes, la majorité des catégories d’équipements de la BPE ont pu être associées à des clés/valeurs d’OSM, à l’exception de quelques-unes d’entre elles (notamment les grandes surfaces commerciales). Ces prétraitements fondés sur l’analyse des ressemblances entre catégories d’équipements sont indispensables pour réaliser l’intégration des bases OSM et BPE et proposer une analyse inédite de la complétude d’OSM à une échelle régionale, pour les communes françaises de la zone d’étude.

2. Analyse de complétude d'OSM à l'échelle régionale (partie française) : comparaison avec la base BPE de l'INSEE

Premier volet de l'analyse de complétude, cette partie repose sur la comparaison systématique des données OSM avec celles de la BPE de l'INSEE, dans la partie française de l'espace d'étude (703 communes) et pour la sélection des 15 équipements d'intérêt (*voir Introduction*). Rappelons qu'il s'agit d'une évaluation relative de la base OSM (Zielstra et Zipf, 2010), les données étant comparées non pas directement à la réalité du terrain mais à un autre jeu de données, qui constitue lui aussi une certaine modélisation des équipements réels : la BPE présente un léger décalage temporel avec OSM (versions du 1er janvier 2018 pour la BPE, de septembre 2019 pour OSM) et offre une représentation des équipements elle-même non dénuée de biais (*voir encadré 1, partie 1*). Cela dit, l'existence d'une base nationale de référence sur les équipements côté français, faisant l'objet d'enquêtes annuelles de qualité, offre une opportunité inédite pour évaluer le degré de complétude des données OSM à une échelle régionale.

L'analyse qui suit s'attache donc à prendre la mesure des écarts entre OSM et la BPE, en comparant aussi bien les quantités d'équipements recensées dans chaque commune que leur diversité, appréhendée à partir de la présence/absence de chaque équipement et des niveaux hiérarchiques qu'ils reflètent. Plusieurs questions générales traversent l'évaluation et l'interprétation de ces écarts entre OSM et la BPE : plus de 10 ans après la création du projet OSM, dans un des pays comptant parmi les communautés OSM les plus actives, le degré de complétude des POI liés aux équipements commerciaux semble-t-il suffisant pour analyser la distribution des équipements commerciaux à partir de cette base contributive ? Observe-t-on, comme l'ont montré la plupart des travaux sur la complétude des réseaux routiers dans OSM (*voir section 1.1*), l'existence d'un gradient décroissant de complétude depuis les communes urbaines les mieux contribuées vers les espaces ruraux isolés ? Et peut-on identifier des catégories d'équipement qui seraient mieux contribuées et donc plus fiables que d'autres dans OSM, sur lesquelles on pourrait s'appuyer pour analyser la répartition spatiale des équipements commerciaux dans les pays voisins, où il n'existe pas de source conventionnelle disponible sur cette thématique ?

2.1. Mesurer les écarts OSM/BPE d'après les quantités communales d'équipements

Des équipements globalement trois fois moins nombreux dans OSM

De manière générale, pour les 15 catégories sélectionnées, **la base OSM recense 2720 équipements** dans la partie française de l'espace d'étude, **contre 8758 dans la BPE**. Il existe donc une nette sous-représentation des équipements dans OSM, par rapport à la BPE. Dans l'ensemble des communes françaises de la zone d'étude, on en dénombre **3,2 fois moins**. Cet écart peut sembler important par rapport à d'autres éléments de la base OSM, notamment aux réseaux routiers, et aux analyses de complétude des réseaux routiers dans différentes zones du nord-ouest de l'Europe (voir partie 1). On peut faire l'hypothèse que les éléments saisis dans la base au titre des POI se prêtent moins bien aux contributions à distance, à partir de l'interprétation d'images satellitaires, « en fauteuil », en dehors de quelques catégories (notamment les stations-services, dont l'emprise spatiale est plus facilement identifiable par

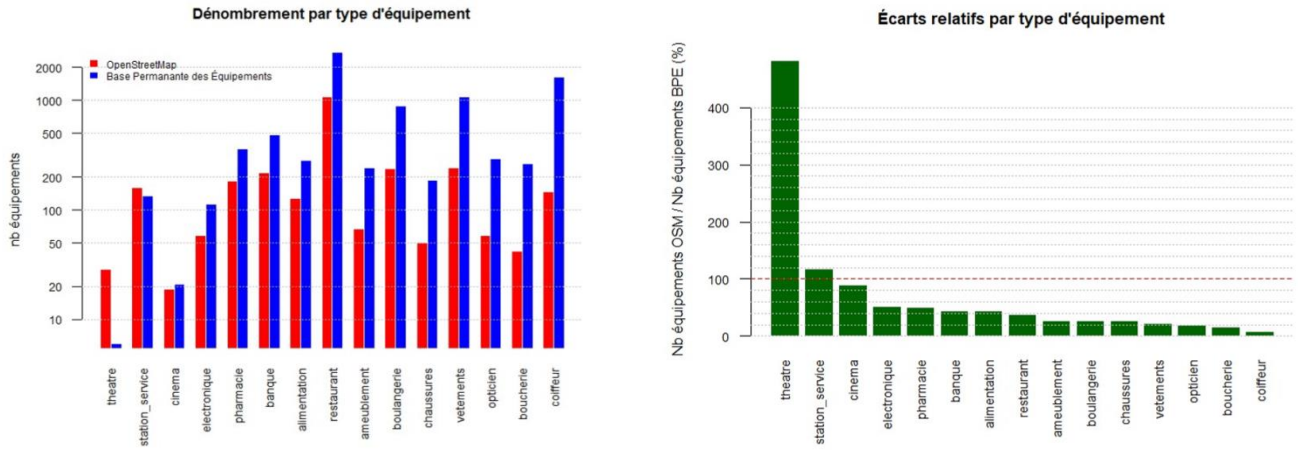
photo-interprétation). Cette catégorie d'objets dépend davantage des observations des contributeurs sur le terrain, même si certains équipements (par exemple les pharmacies) peuvent être renseignés à l'aide de bases professionnelles importées dans OSM.

Des écarts variables selon les types d'équipements

Malgré cette sous-estimation globale, la complétude des données OSM est très variable selon les types d'équipements. On peut distinguer quatre grands groupes à la lumière des écarts observés avec la BPE (*Figure 9*) :

- **Groupe 1 :** pour la moitié des catégories (7 sur 15), les équipements sont **fortement sous-représentés dans OSM, 4 à 10 fois moins** nombreux que dans la BPE : il s'agit des catégories **boucherie, boulangerie, coiffeur** (gamme de proximité dans la BPE), **ameublement, chaussures, opticien, vêtement** (gamme intermédiaire).
- **Groupe 2 :** Pour 5 catégories, la sous-représentation est plus modérée, les équipements sont **2 à 3 fois moins** nombreux dans OSM : **alimentation, café-restaurant, pharmacie** (gamme de proximité dans la BPE), **banque, électronique** (gamme intermédiaire).
- **Groupe 3 :** Pour deux catégories, les **stations-services** et les **cinémas**, les **effectifs sont proches** de ceux de la BPE (respectivement +19% et -10% d'équipements dans OSM par rapport à la BPE). Dans le cas des cinémas, le faible effectif de ces équipements de la gamme supérieure ainsi que leur caractère emblématique dans l'équipement culturel des communes peuvent jouer pour expliquer cette bonne représentation. Le cas des stations-services est particulier et plusieurs hypothèses peuvent être avancées pour expliquer leur sur-représentation non négligeable dans OSM : le critère restrictif fixé pour l'enregistrement de l'équipement dans la BPE, en fonction d'un seuil de volume d'essence servi (encadré 1), la plus grande facilité d'identification de l'emprise spatiale de l'équipement à partir de photo-interprétations, ainsi que l'intérêt de l'équipement pour les services de calculs d'itinéraires proposés à partir de la base OSM.
- **Groupe 4 :** les équipements nettement plus nombreux dans OSM. Cela correspond au cas particulier des **théâtres**, environ 5 fois plus nombreux dans OSM, du fait d'une définition beaucoup plus large que dans la BPE, où seules figurent les scènes nationales.

Figure 9 – Nombre d'équipements et écarts OSM-BPE, par catégorie d'équipement



Des écarts variables selon les types d'espaces

La cartographie des écarts OSM-BPE (*Figure 10*) amène par ailleurs à s'interroger sur la variation spatiale de la complétude d'OSM. Les quelques travaux consacrés à la question de l'hétérogénéité spatiale de la base OSM à une échelle régionale ont mis en évidence l'existence de gradients de contributions autour des centres de grandes métropoles (voir partie 1.1), avec une contribution décroissante à mesure que l'on s'éloigne du centre. Dans les cartes de la *Figure 10*, la distribution spatiale des écarts ne semble pas aussi claire à première vue, les zones de plus forte sous-représentation relative des équipements dans OSM se situant aussi bien dans des espaces ruraux de faible densité que dans des zones densément urbanisées (au nord du Sillon lorrain, notamment). Non seulement les équipements semblent plus concentrés dans un petit nombre de communes dans OSM (par exemple, la commune de Metz concentre 34% des équipements de proximité dans OSM, contre 17% dans la BPE) mais aussi les hiérarchies de communes par effectifs d'équipements ne sont pas tout à fait les mêmes d'une base à l'autre (Metz et Thionville restent les 2 premières communes les plus équipées, mais dans OSM les 3 suivantes par ordre d'importance sont Forbach, Sarreguemines et Longwy, tandis que dans la BPE il s'agit de Sarreguemines, Forbach et Saint-Avold) (voir site Web, section « BPE-OSM (France)).

Afin de préciser le lien éventuel entre la complétude d'OSM et l'appartenance des communes à différents types d'espaces, nous sommes partis pour identifier ces types d'espaces d'une classification en cinq grandes catégories de communes, qui s'appuie à la fois sur le zonage en aires urbaines et sur la composition communale des unités urbaines de l'INSEE : sont ainsi distinguées trois catégories de communes au sein des aires urbaines (centre, banlieue, périurbain) et deux catégories en dehors des aires urbaines (ville isolée - correspondant aux plus petites agglomérations, par exemple Boulay-Moselle à l'est de Metz -, commune rurale hors périurbain).

Figure 10 – Répartition communale des équipements et variation spatiale des écarts OSM-BPE (exemple des équipements de la gamme de proximité)

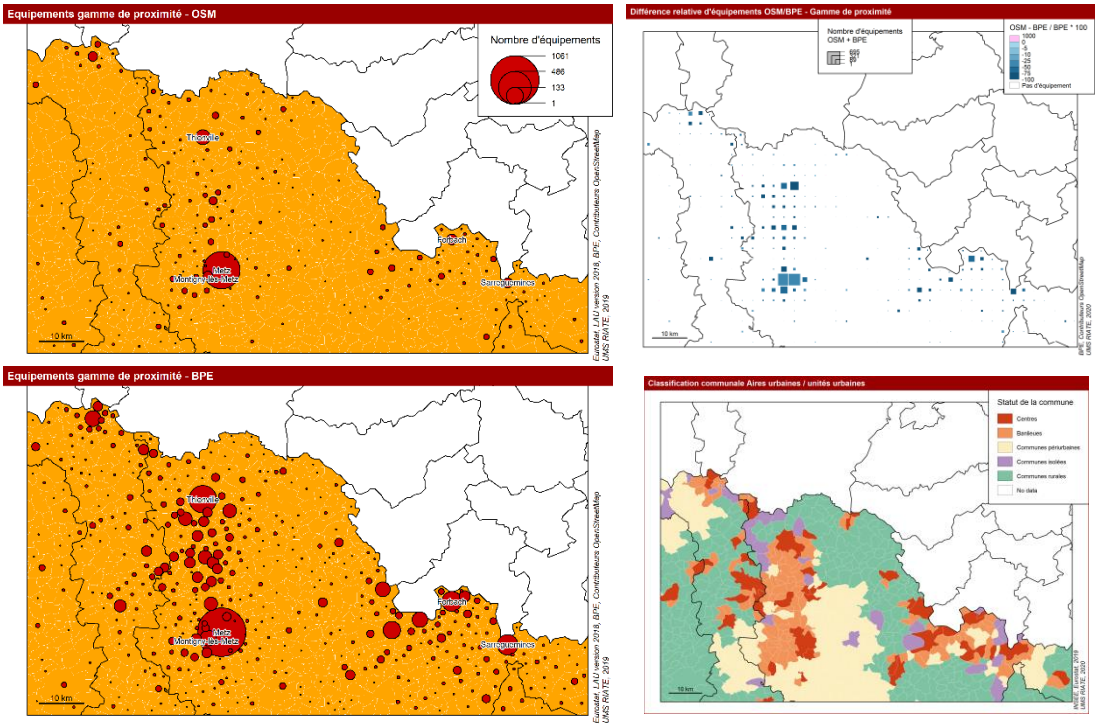
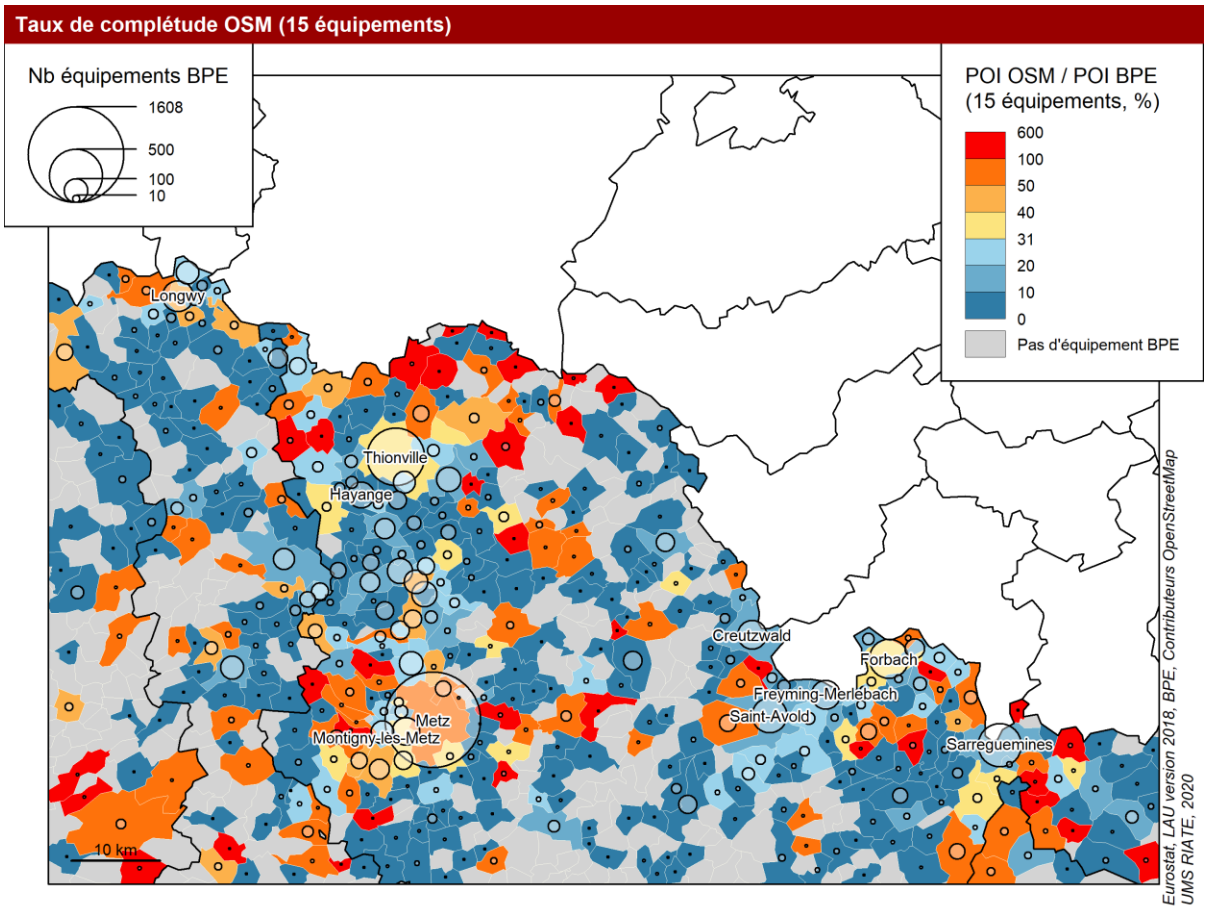
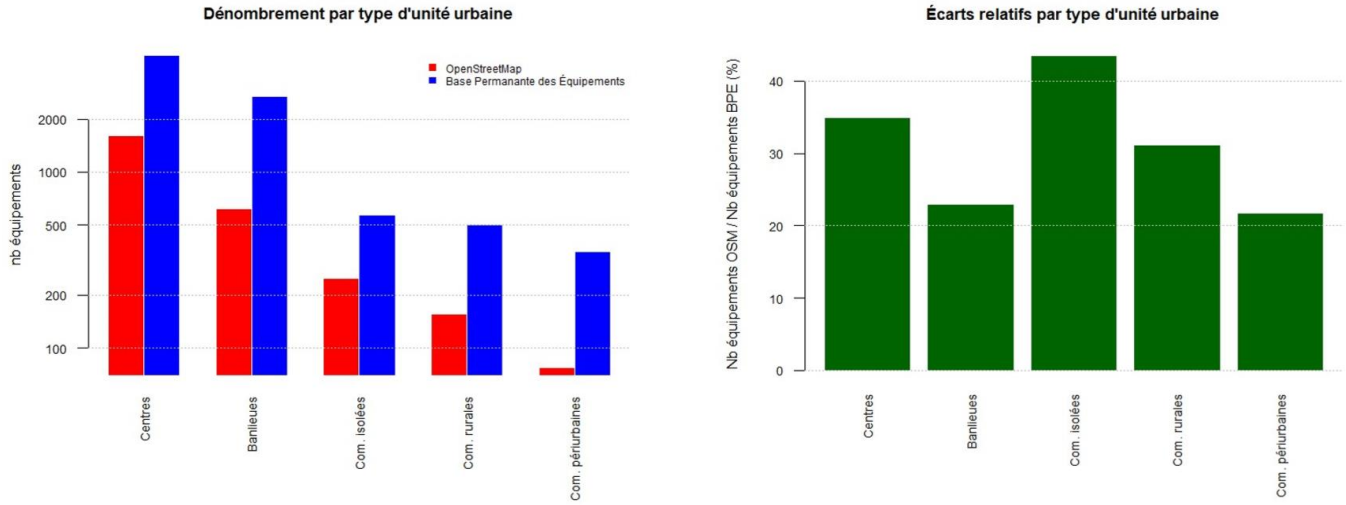


Figure 11 – Taux de complétude d'OSM dans les communes françaises de l'espace d'étude, par rapport à la BPE (pour 15 équipements)



L'analyse des écarts relatifs d'équipements OSM-BPE par catégorie de commune (*Figure 12*) montre que la **complétude d'OSM est plus élevée dans les communes urbaines isolées** (environ deux fois moins d'équipements dans OSM, le nombre d'équipements dans OSM représentant 44% du nombre d'équipements dans la BPE) ainsi que **dans les communes centres d'aires urbaines** (environ trois fois moins d'équipements dans OSM). Les communes **les moins bien renseignées dans OSM** par rapport à la BPE (environ cinq fois moins d'équipements) sont les **communes périurbaines et les communes de banlieue**. La relation entre le degré de complétude et le type d'espace n'est toutefois pas si claire, dans la mesure où en moyenne, les communes rurales hors aires urbaines sont nettement mieux renseignées que les communes périurbaines et de banlieue (environ trois fois moins d'équipements dans OSM, soit le même ordre de grandeur que les communes centres d'aires urbaines). **Ce qui ressort le plus, à l'échelle régionale, est donc la sous-représentation des équipements d'OSM dans les communes de banlieue et du périurbain.**

Figure 12 – Nombre d'équipements et écarts OSM-BPE, par catégorie de commune

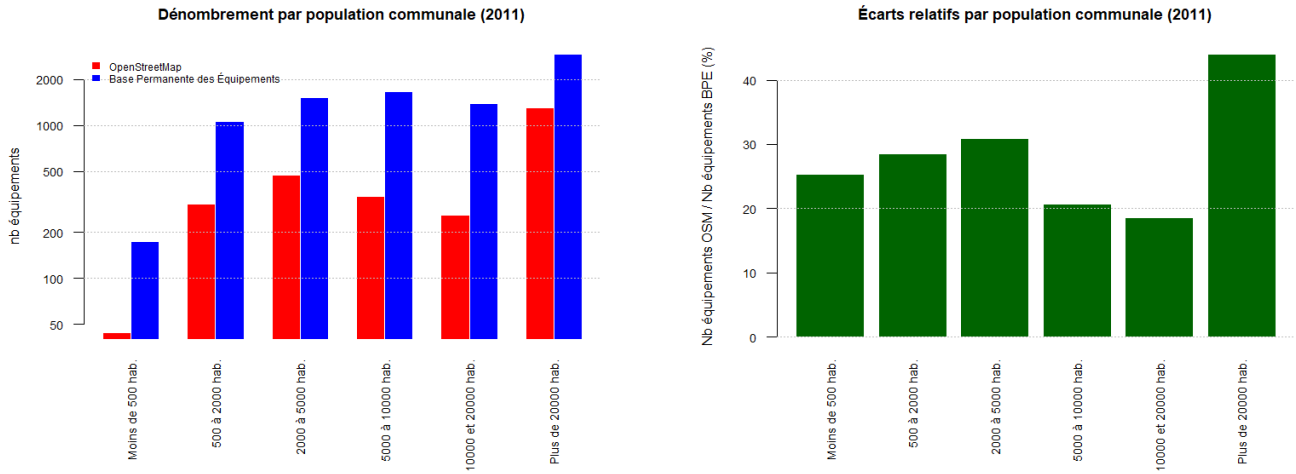


Categorie	Nb de communes	OSM	BPE	Ecart Absolu	Ecart relatif
Centres	44	1621	4639	-3018	34.9
Banlieues	118	618	2696	-2078	22.9
Com. isolées	27	248	569	-321	43.6
Com. rurales	312	156	500	-344	31.2
Com. périurbaines	202	77	354	-277	21.8

Ces résultats peuvent être complétés et précisés en analysant le lien entre le degré de complétude d'OSM et la taille des communes, partant de l'hypothèse selon laquelle les communes les plus peuplées se caractériseraient par une plus grande exhaustivité de la base OSM (*Figure 13*). Cette hypothèse est là encore partiellement validée : les données OSM sont en effet plus complètes dans les communes de plus de 20 000 habitants, où elles représentent 44% du nombre d'équipements BPE (55% dans la seule commune de Metz centre). Et cette complétude tend à augmenter avec la taille dans les communes jusqu'à 5 000 habitants (environ trois à quatre fois moins d'équipements dans OSM par rapport à la BPE). **Mais les**

communes de 5 000 à 20 000 habitants présentent une complétude bien plus faible (cinq fois moins de données dans OSM) qui amène à relativiser le lien entre taille démographique et degré de complétude.

Figure 13 – Nombre d'équipements et écarts OSM-BPE, par classe de taille des communes



Categorie	Nb de communes	OSM	BPE	Ecart Absolu	Ecart relatif
Moins de 500 hab.	326	44	174	-130	25.3
500 à 2000 hab.	252	305	1067	-762	28.6
2000 à 5000 hab.	79	472	1524	-1052	31.0
5000 à 10000 hab.	28	344	1666	-1322	20.6
10000 et 20000 hab.	13	259	1389	-1130	18.6
Plus de 20000 hab.	5	1296	2938	-1642	44.1

Ces premières analyses des écarts OSM/BPE, basées sur la comparaison de la quantité d'équipements recensés dans les deux sources, ont permis d'évaluer précisément le degré de sous-estimation ou plus rarement de surestimation des équipements dans OSM par rapport à la BPE, tout en mettant en exergue l'hétérogénéité spatiale et thématique des équipements les mieux et les moins bien contribués. On a pu ainsi identifier les équipements les mieux contribués dans OSM (théâtres et cinémas, stations-services) et ceux pour lesquels les lacunes étaient relativement moins importantes, parmi la gamme des équipements de proximité (cafés-restaurants, pharmacies, alimentations) comme des équipements intermédiaires (banques, magasins d'électronique). D'autre part, si dans les communes-centres des villes les données OSM sont plus complètes qu'ailleurs, ce sont les communes de banlieue et du périurbain (et non pas les communes rurales hors aires urbaines) qui ressortent comme les moins bien renseignées dans OSM.

2.2. Mesurer les écarts OSM/BPE d'après la diversité des équipements présents dans les communes

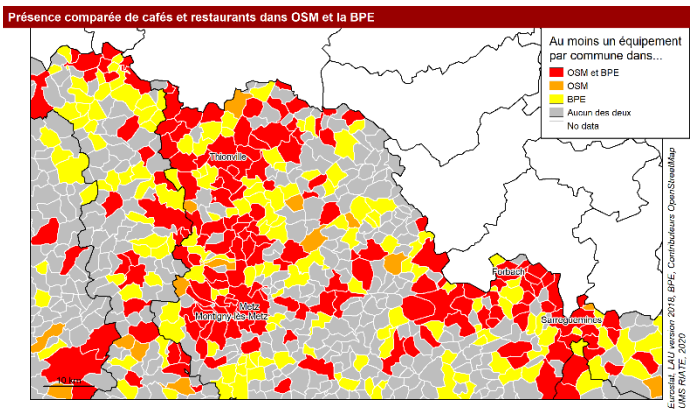
Cette première comparaison peut être complétée par une analyse de la diversité des équipements présents dans chaque commune, qui fait abstraction de leur masse globale. Le fait de tenir compte de la présence d'un type d'équipement plutôt que de sa quantité est une autre approche de la complétude d'OSM, qui devrait *a priori* minimiser les écarts de mesure du phénomène observé entre les deux bases.

Présence/absence d'équipements dans les communes

L'observation de la présence des 15 équipements dans les 703 communes de la zone d'étude permet d'affiner les constats dressés précédemment à partir des dénombrements globaux. En nombre de communes équipées, les écarts sont un peu plus faibles que lorsque l'on considère les quantités d'équipements mais restent importants. Globalement, il y a presque **deux fois plus de communes équipées (comptant au moins 1 équipement parmi les 15 catégories observées) dans la BPE** (452 contre 233 dans OSM).

Dans le détail des catégories, on retrouve des différences qui rappellent celles qui ont été mesurées à partir des effectifs d'équipements. A titre d'exemple, on peut illustrer les résultats obtenus pour quatre de ces équipements (*Figure 14*, *Tableau 1*). Pour les cafés-restaurants, il y a 2,2 fois moins de communes équipées dans OSM que dans la BPE (147 contre 321). Pour les banques, l'écart est equivalent, avec 143 communes équipées dans la BPE et 72 communes équipées dans OSM. L'écart est moindre pour les pharmacies (83 communes équipées dans OSM contre 148 dans la BPE) et encore plus faible pour les cinémas (15 communes équipées dans OSM, 19 dans la BPE). Par ailleurs, on notera que **les communes où un équipement est présent dans la BPE mais pas dans OSM ne se réduisent pas aux communes rurales** : il s'agit parfois de communes urbaines de plus de 10 000 habitants, comme dans la commune-centre de Hayange, qui compte 37 cafés-restaurants, 5 pharmacies et 9 banques d'après la BPE et où aucun de ces équipements n'est recensé dans OSM. On peut citer aussi les cas de Sarreguemines (14 banques dans la BPE, aucune dans OSM), de Saint-Avold et de Woippy (cinema recensé dans la BPE, pas dans OSM).

Figure 14 – Présence comparée d'équipements dans OSM et la BPE (cafés-restaurants et banques)



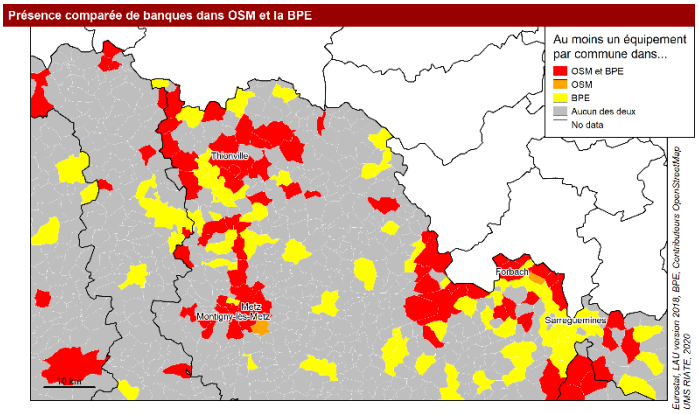


Tableau 1 - Nombre d'équipements présents d'après OSM et la BPE (15 équipements de l'étude)

Equipements	Nb communes non équipées (OSM – BPE)	Nb communes équipées dans OSM et/ou BPE	dont … BPE : oui OSM : oui	dont … BPE : oui OSM : non	dont … OSM : oui BPE : non
Théâtre	**687**	**16**	4	1	11
Cinéma	**682**	**21**	13	6	2
Chaussure	**651**	**52**	15	37	0
Electronique	**637**	**66**	13	45	8
Ameublement	**627**	**76**	25	47	4
Opticien	**622**	**81**	23	58	0
Vêtement	**605**	**98**	29	69	0
Station-serv.	**604**	**99**	80	1	18
Boucherie	**585**	**118**	22	93	3
Alimentation	**576**	**127**	29	76	22
Banque	**558**	**145**	70	73	2
Pharmacie	**555**	**148**	83	65	0
Boulangerie	**447**	**256**	98	157	1
Coiffeur	**399**	**304**	52	252	0
Café-restaur.	**367**	**336**	164	157	15
Tout confondu (15 équipements)	**239**	**464**	220	232	12

Diversité des équipements dans les communes

Plus globalement, cette comparaison peut être synthétisée à l'aide d'une première mesure simple de diversité, qui s'appuie sur le nombre d'équipements différents par commune (de 0 à 15) et sur une discrétisation en quatre niveaux de cette valeur (aucun équipement, faible diversité de 1 à 4, moyenne de 5 à 9, forte de 10 à 15). Les résultats (*Figure 15*, *Tableau 2*) concordent entre OSM et BPE pour un peu plus de la moitié des communes : 369 d'entre elles sont classées dans la même catégorie. 12 d'entre elles, en particulier, ressortent comme très diversifiées dans les deux bases (dont Metz, Thionville, Sarreguemines, Forbach, Longwy, Montigny-lès-Metz). Toutefois, les cas de non correspondance sont aussi nombreux avec, une majorité de cas pour lesquels la diversité d'équipements est moindre dans OSM : **d'après OSM, il y a par exemple cinq fois moins de communes très diversifiées que dans la BPE et quasiment deux fois plus de communes non équipées**. Dans le détail, on constate à nouveau l'existence d'écarts très importants dans quelques communes assez peuplées. Par exemple, parmi les 57 communes qui présentent une forte diversité d'équipements dans la BPE, 19 n'ont qu'une faible diversité dans OSM (c'est le cas à Hayange, Stiring-Wendel, Villerupt) et 2 n'enregistrent aucun des équipements observés dans l'étude (cas d'Auboué et de Marange-Silvange, en périphérie de Metz).

Figure 15 – Nombre d'équipements différents présents dans les communes de l'espace d'étude

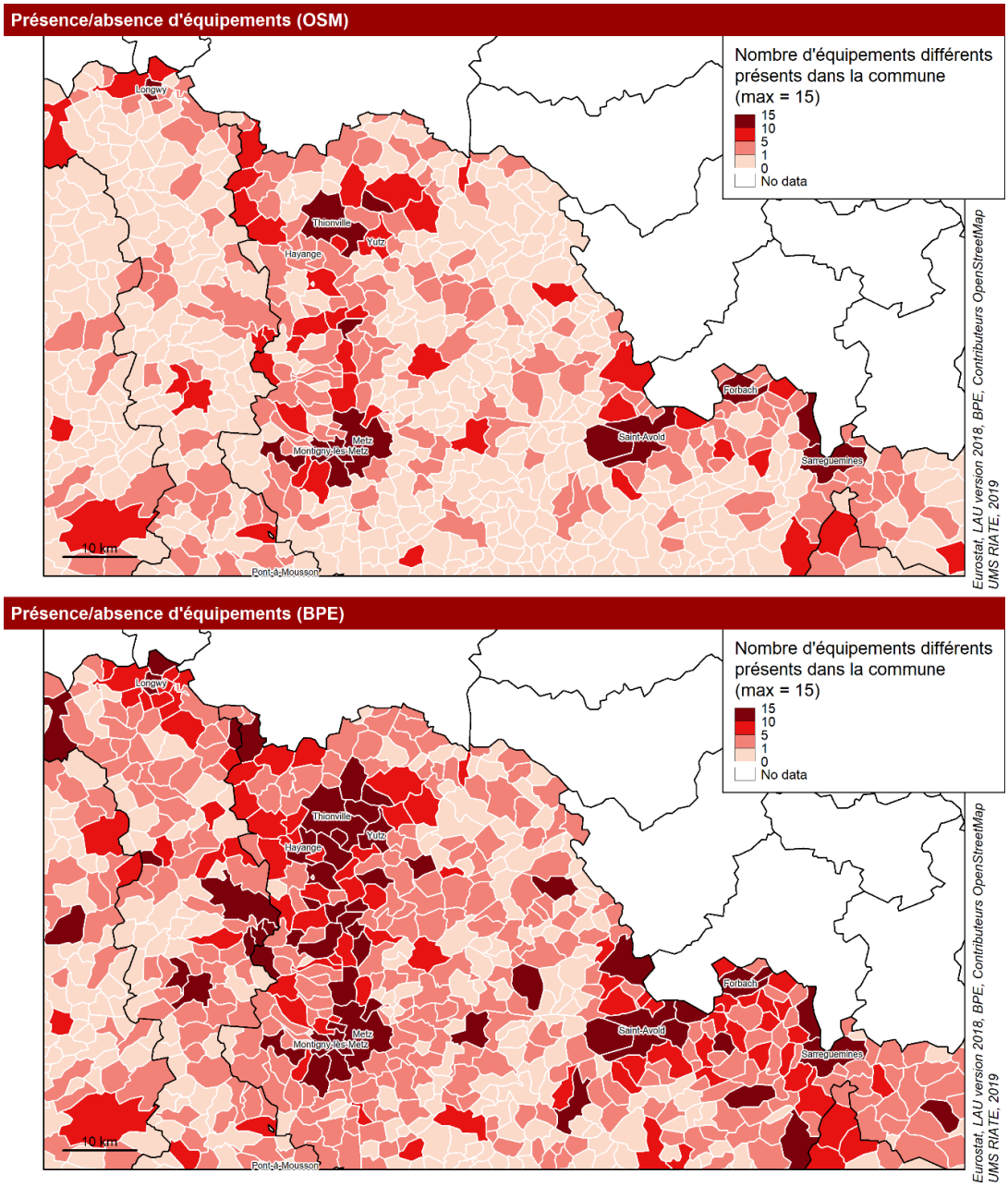


Tableau 2 - Diversité des équipements présents dans les communes de l'espace d'étude

BPE/OSM	Pas équipée	Faible	Moyenne	Forte	Total
Pas équipée	**243**	8			251
Faible	210	**94**	1		305
Moyenne	15	55	**20**		90
Forte	2	19	24	**12**	57
Total	470	176	45	12	**703**

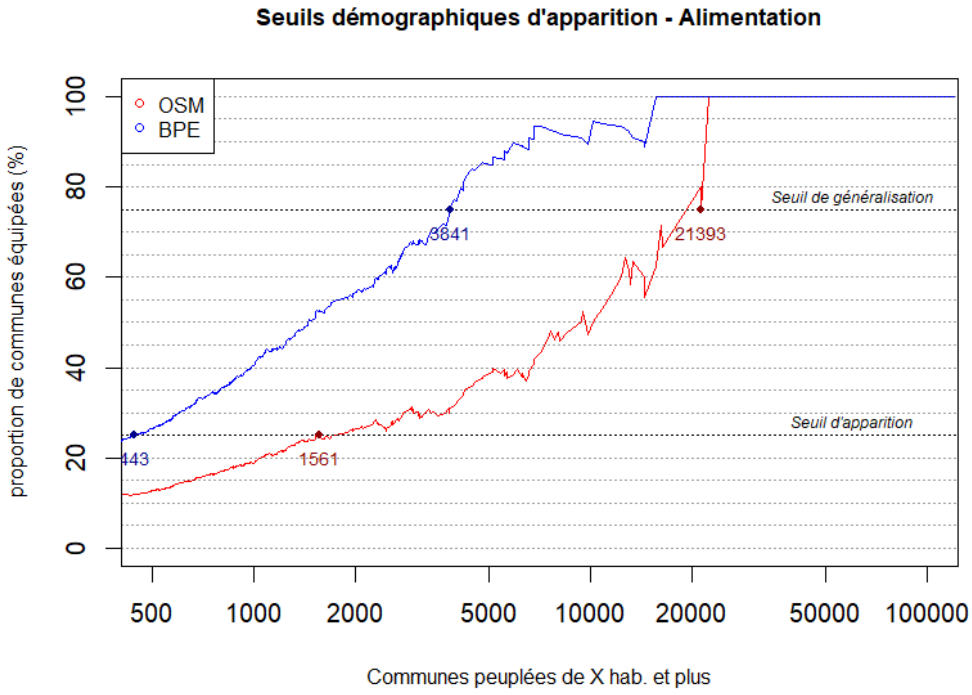
NB : Dans le tableau ci-dessus, la diversité des équipements est évaluée par le nombre d'équipements différents présents dans chaque commune, pour les 15 équipements de l'étude : elle est faible de 1 à 4 équipements, moyenne de 5 à 9, forte de 10 à 15.

Equipements et population communale : seuils d'apparition et de généralisation

Dans la plupart des analyses consacrées à l'encadrement des territoires par les commerces et les services, l'information sur la présence d'équipements est souvent croisée avec celle de la population des communes équipées afin d'identifier des hiérarchies d'équipements : équipements fréquents parmi les communes peu peuplées pour les plus ubiquistes d'entre eux, concentrés dans les communes les plus peuplées pour les plus rares d'entre eux. On s'appuie ici sur des indicateurs de seuils d'apparition et de généralisation[11] souvent mobilisés dans l'analyse des hiérarchies d'équipements (voir par exemple, Barzak et Hilal, 2007, Baudet et al. 2019) pour comparer OSM et la BPE au regard des hiérarchies d'équipements que les données des deux bases font ressortir. Nous faisons l'hypothèse que malgré la sous-représentation globale des équipements dans OSM, les mêmes hiérarchies relatives d'équipement ressortent d'OSM et de la BPE.

Un premier exemple de lecture et d'interprétation de ces seuils peut être donné pour la catégorie Alimentation (*Figure 16*) : le seuil d'apparition des magasins d'alimentation est de 1 500 habitants dans OSM (un quart des communes de moins de 1500 hab. comptent un magasin d'alimentation), tandis que ce seuil est trois fois plus bas dans la BPE (440 hab.). Quant au seuil de généralisation, il est de 21 000 habitants dans OSM (l'équipement est implanté dans trois quarts des communes ayant moins de 21 000 habitants) et de 3 800 habitants dans la BPE.

Figure 16 – Seuils d'apparition et de généralisation de l'équipement alimentation



[11] « Le seuil d'apparition mesure le nombre d'habitants à partir duquel 25% des communes possèdent le service. C'est-à-dire que 25% des communes équipées ont une population inférieure ou égale à cette taille. Le seuil de généralisation mesure le nombre d'habitants à partir duquel 75% des communes possèdent le service. C'est-à-dire que 75% des communes équipées ont une population inférieure ou égale à cette taille » (Baudet et al., 2019). Pour obtenir ces seuils, on trie les communes de l'espace d'étude par ordre croissant de population, ces communes étant décrites par les présences et absences d'équipements (codées 1/0). Ensuite et pour chaque commune, on calcule la part des communes plus peuplées (incluant cette commune) qui disposent de cet équipement (fréquences cumulées).

Plus généralement, la prise en compte de l'un des seuils (par exemple le seuil d'apparition) et la comparaison de ce seuil pour l'ensemble des équipements permet de voir dans quelle mesure le choix de la source BPE ou OSM bouscule la représentation de la hiérarchie des équipements. La *Figure 17* qui compare les seuils d'apparition dans la BPE et dans OSM pour l'ensemble des équipements montre que ces seuils sont globalement plus bas dans la BPE que dans OSM, à l'image de ce que l'on vient de voir pour l'alimentation, à l'exception des cafés-restaurants, des stations-service et des théâtres. Dans le *Tableau 3*, la reconstitution de gammes d'équipements à partir des valeurs similaires de seuils d'apparition montre que **8 équipements classent les communes de la même manière, quelle que soit la source : cafés-restaurants (grande proximité), banques et pharmacies (proximité), ameublement, électronique, opticiens et vêtements (intermédiaire), cinéma (supérieur)**. Certains équipements, à l'inverse, apparaissent plus ubiquistes dans OSM que dans la BPE (les restaurants, les stations-services, les théâtres), tandis que d'autres semblent plus sélectifs dans OSM (boulangeries, coiffeur, boucherie, chaussures).

Figure 17 – Seuils d'apparition pour les équipements, comparaison OSM-BPE

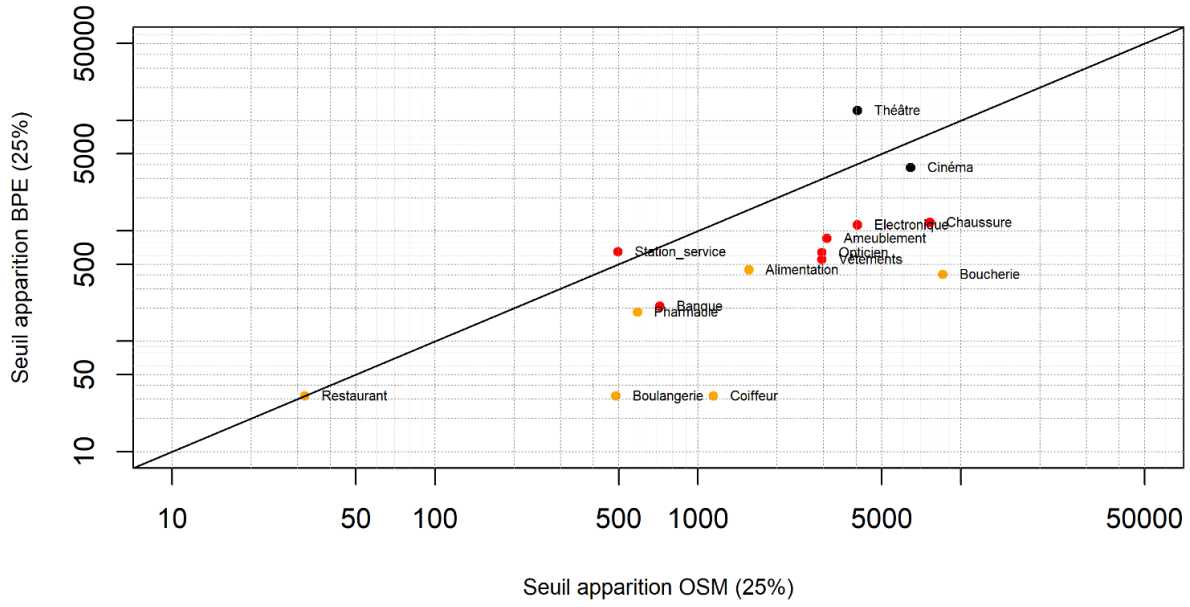


Tableau 3 - Reconstitution de gammes d'équipements à partir des seuils d'apparition dans OSM/BPE

OSM / BPE	***Grande proximité*** *(30 hab.)*	***Proximité*** *(200 - 1600 hab.)*	***Intermédiaire*** *(2900 - 4000 hab.)*	***Supérieure*** *(6400 - 8500 hab.)*
Grande proximité *(30 hab.)*	**Café-restaurant**	Boulangerie Coiffeur		
Proximité *(200 - 300 hab.)*		**Pharmacie** **Banque**		
Intermédiaire *(400 - 1200 hab.)*		Station-service Alimentation	**Ameublement** **Electronique** **Opticien Vêtements**	Boucherie Chaussures
Supérieure *(3700 - 12400 hab.)*			Théâtre	**Cinéma**

Typologie communale : profils d'équipement

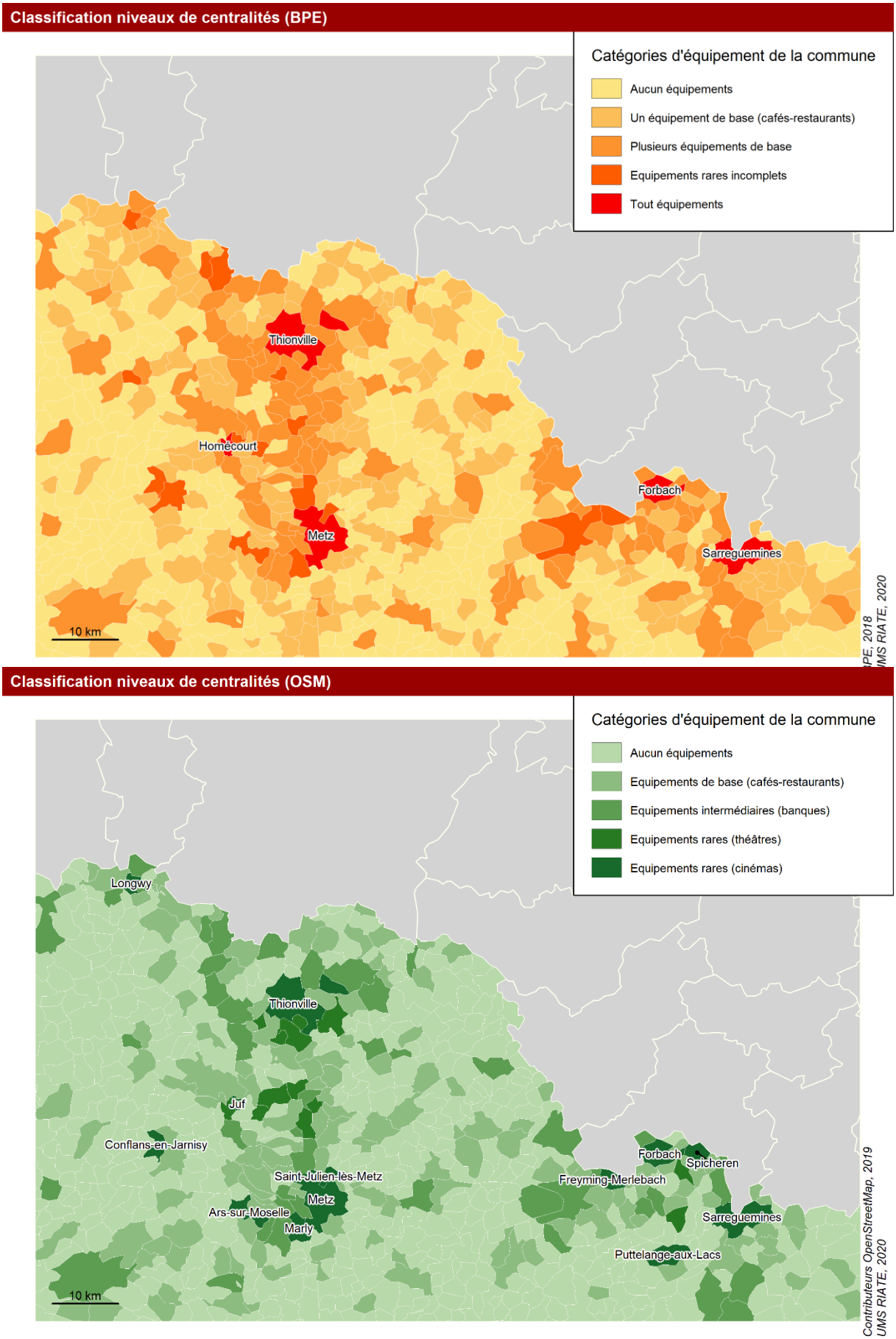
Une **synthèse des observations conduites sur les présences et les seuils d'apparition des équipements peut être réalisée à l'aide de méthodes d'analyses multivariées**. Il s'agit d'identifier les principales corrélations entre équipements présents, puis de regrouper les communes caractérisées par les mêmes combinaisons d'équipements. Autrement dit, on cherche à classer les communes sur la base des combinaisons de modalités de présence-absence (0/1) d'équipements qu'elles ont en commun (Husson et al, 2016) : elles seront regroupées si elles ont en commun un grand nombre de modalités (les mêmes équipements sont présents ou sont absents) ou bien si elles ont en commun une modalité rare (par exemple la présence d'un cinéma), quelles que soient leurs différences par ailleurs.

Deux méthodes d'analyse multivariée ont été successivement appliquées (partie 4.4 de la partie du site Web dédiée à la comparaison BPE-OSM) : une Analyse en Composantes Multiples (ACM) est d'abord réalisée afin de transformer les variables qualitatives (modalités de présence ou d'absence d'équipement) en variables continues, ces nouvelles valeurs étant ensuite traitées à l'aide d'une Classification Ascendante Hiérarchique (CAH) pour faire ressortir les principaux profils de communes. Les analyses sont menées à partir d'une sélection restreinte de 6 équipements, qui sont les équipements les mieux renseignés dans OSM d'après l'analyse de complétude : les cafés-restaurants et les pharmacies pour les équipements les plus fréquents, les banques, les magasins d'électronique et les théâtres à un niveau intermédiaire, les cinémas en tant qu'équipement plus rare. On retient une classification en 5 catégories de communes, dont le profil varie sensiblement selon que l'on s'appuie sur la base OSM ou sur la BPE (*Tableau 4*).

Tableau 4 - Niveaux d'équipements présents dans les communes françaises de la zone d'étude, pour une sélection de 7 équipements

Niveaux	BPE	OSM
1	**Aucun équipement (379 communes)**	**Aucun équipement (506 communes)**
2	**Café-restaurant (195 communes)**	**Café-restaurant, parfois pharmacie et station-service (116 communes)**
3	**Café-restaurant, Pharmacie, Banque (195 communes)** *Ex : Hayange, Florange, Fameck... mais aussi communes de banlieue de Metz*	**Café-restaurant, Pharmacie, Banque (52 communes)** *Ex :*
4	**Cinéma, parfois magasin d'électronique (15 communes)** *Ex : Longwy, Amnéville, St-Avold, Woippy (tout sauf un théatre)... mais aussi qqs communes de 2000 habitants (Conflans en Jarnisy, banlieue de Jarny, possède tous équipements -y compris cinéma- sauf banque et théatre. Même chose pour St-Julien-les-Metz, Sérémange-Erzange en banlieue de Hayange)*	**Théâtre, pas de cinéma (14 communes)**
5	**Cinéma et théâtre (5 communes)** *Ex : Metz, Thionville, Forbach, Sarreguemines... mais aussi Homécourt (6000 habitants, appartient à l'agglo de Joeuf, à 20 km au N-O de Metz), parce que la commune possède un théâtre (la Machinerie, scène conventionnée d'intérêt national)*	**Cinéma, pas de théâtre (15 communes)** *Ex : Metz, Sarreguemines, Thionville, Forbach, Longwy + périphérie de Metz (St Julien...banque, cinéma, restaurant)*

Figure 18 – Localisation des communes en fonction de leur profil d'équipement



La comparaison des profils communaux d'équipements selon la source choisie montre globalement que :

- On obtient des combinaisons d'équipements similaires, sauf pour les niveaux supérieurs de la hiérarchie où selon les cas, c'est la présence des cinémas (dans OSM) ou des théâtres (dans la BPE) qui s'avère la plus sélective.
- Sans surprise, il y a beaucoup plus de communes décrites comme « non équipées » à partir des données OSM (506 contre 379 communes)
- Il y a une quinzaine de communes de type dit « intermédiaire » d'après la BPE (présence d'au moins une banque, pharmacie…) qui sont classées parmi les communes non équipées dans OSM : ce sont des communes de 2000 habitants en moyenne, en majorité des communes rurales multipolarisées, mais aussi des communes de banlieue (de Longwy, de Metz, de Sarreguemines…).
- Inversement, d'après OSM, le choix des théâtres et des cinémas fait entrer dans le niveau supérieur des communes de faible poids démographique. Les théâtres sont un marqueur plus sélectif du niveau supérieur des centralités dans la BPE. Difficile de trouver un marqueur de ce niveau supérieur dans OSM.

Conclusion partie 2

Dans l'ensemble des communes françaises de la zone d'étude, les **équipements OSM** correspondant aux 15 équipements sélectionnés sont **environ 3 fois moins nombreux que dans la BPE de l'INSEE** et il y a 2 fois plus de communes non équipées (aucun des 15 équipements) dans OSM. Au-delà de cette comparaison globale, la **complétude est très variable selon les contextes spatiaux** : plus élevée dans les communes de centres-villes, que ce soit au sein d'aires urbaines ou d'agglomérations isolées, plus faible dans les communes rurales hors zone urbaine et surtout dans les banlieues et le périurbain. La **variation de la complétude suit donc une logique spatiale plus complexe qu'un simple gradient centre-périphérie**. C'est ce que montrent aussi les résultats sur la présence/absence des équipements, certaines zones urbaines semblant largement sous-contribuées, malgré leurs fortes densités de population et d'équipements (voir les communes-centres de Hayange et Sarreguemines). Par ailleurs, les cas de figure sont **très contrastés d'un équipement à l'autre**. Les équipements les plus complets dans OSM correspondent à ceux dont la définition est moins restrictive dans OSM que dans la BPE : c'est le cas surtout des théâtres et, à un moindre degré, des stations-services. Sur les 13 types d'équipements restants, 6 présentent une bonne complétude relative, comprise entre 40% et 90% : il s'agit des cafés-restaurants, magasins d'alimentation, banques, pharmacies, magasins d'électronique et cinémas (par ordre croissant de complétude).

En partant de la sélection des 6 équipements les plus fiables d'après les deux volets de l'analyse de complétude (cafés-restaurants pour les équipements les plus fréquents, pharmacies et banques pour les équipements de proximité, magasins d'électronique pour les équipements intermédiaires, théâtres et cinémas pour les équipements de niveau supérieur), nous avons fait ressortir **5 profils communaux d'équipements** qui reflètent différents niveaux hiérarchiques d'équipement et présentent de **nombreuses similarités selon qu'ils sont construits à partir d' OSM ou de la BPE**. La classification des communes d'après ces profils montre plusieurs traits communs entre OSM et la BPE, **70% des communes étant associées au même profil d'équipement**. Toutefois, elle souligne également que les données OSM conduisent à sous-estimer fortement le bas de la hiérarchie des communes équipées, tout en sous-estimant la diversité des équipements dans quelques villes (par exemple à Hayange).

3. Analyse de complétude d'OSM à l'échelle locale : comparaison avec des relevés de terrain dans 6 zones tests (France, Allemagne)

Après avoir considéré l'ensemble des communes françaises de la zone d'étude, ce second volet de l'analyse de complétude se focalise sur **quelques zones-tests en France et en Allemagne** et s'appuie sur une **expertise locale et des relevés d'équipements effectués sur le terrain**. Cette analyse complémentaire doit permettre à la fois de préciser la nature des lacunes constatées à l'échelle régionale côté français et d'élargir la comparaison à plusieurs localités allemandes proches de la frontière. Dans la partie française, nous avons vu que la BPE, source institutionnelle, ne pouvait être considérée comme l'unique base de référence pour qualifier la sur ou sous-représentation des équipements dans OSM, compte tenu des biais éventuels pouvant affecter le recensement des équipements dans la base de l'INSEE. Les relevés de terrain permettront dans ce cas de vérifier la validité des constats dressés à partir de la comparaison OSM-BPE. Du côté allemand, l'enquête de terrain vient compenser ponctuellement l'absence d'une base institutionnelle équivalente à la BPE de l'INSEE. En sondant les différences de complétude entre la France et l'Allemagne, il s'agit d'apporter quelques éléments d'analyse sur les différences de contribution de part et d'autre de la frontière, éléments partiels compte tenu du nombre restreint de sites[12] mais d'une grande précision sur un problème généralement peu documenté.

Cette partie s'attache donc à **comparer les données OSM aux relevés de terrain effectués à une échelle infra-communale, pour une sélection de rues commerçantes situées dans six localités (quatre en France, deux en Allemagne)**. Nous présentons d'abord les critères de sélection des terrains et le protocole d'enquête retenu, avant de voir en quoi ces analyses locales permettent de mieux qualifier et de mieux comprendre la complétude de l'information dans OSM, ainsi que sa variabilité selon les contextes spatiaux et les catégories d'équipement. La démarche méthodologique ainsi que l'ensemble des résultats sont détaillés dans la partie 3 du site Web, dédiée aux analyses de terrain.

3.1. Choix des sites d'enquête, collecte et mise en forme des relevés de terrain

Choix des sites d'enquête

Les relevés de terrain ont été réalisés entre septembre 2019 et février 2020 par l'équipe du LOTERR[13] (Université de Lorraine), qui travaille depuis plusieurs années sur l'équipement commercial dans la Grande région. Le choix des zones-tests répond à plusieurs critères pertinents au regard de la thématique d'analyse. La sélection des localités vise d'une part à représenter différents niveaux de la hiérarchie urbaine, depuis le niveau métropolitain jusqu'à celui des petites villes. Il s'agissait en outre de prendre en compte plusieurs contextes

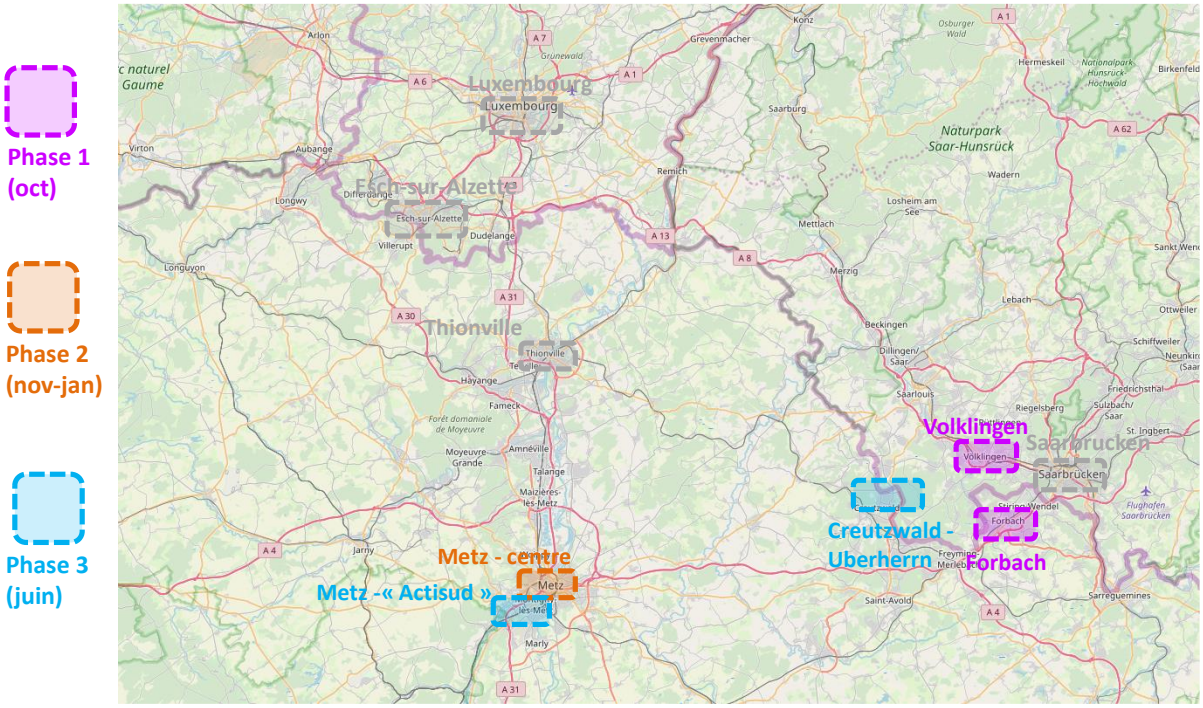
[12] Initialement, l'enquête devait se déployer sur un nombre plus élevé de sites en Allemagne et au Luxembourg ; compte tenu des contraintes de déplacement imposées par le confinement au printemps 2020, seuls deux sites allemands ont pu être conservés.

[13] La grande majorité des relevés ont été réalisés spécifiquement pour ce projet pendant entre janvier 2020 et juillet 2020. Seuls les relevés de Forbach et Völklingen ont été réalisés en septembre 2019 par F. Morel-Doridat dans le cadre de sa thèse de doctorat effectuée au LOTERR.

nationaux, afin d'appréhender la variabilité internationale des contributions. Le choix des sites en Allemagne est guidé par l'intérêt que présente la proximité de villes voisines de la frontière franco-allemande : les villes de Forbach (FR) et Völklingen (DE) d'une part, de Creuzwald (FR) et Überherrn (DE) d'autre part, ne sont distantes que de quelques kilomètres. A l'échelle infra-urbaine, il s'agissait enfin de sonder plusieurs types de centralités commerciales, depuis les rues commerçantes très spécialisées du centre historique d'une métropole régionale, jusqu'aux zones commerciales de périphérie, en passant par les centres de petites villes. Le nombre de sites étant limité par la durée de l'étude et le coût des enquêtes, nous nous sommes arrêtés sur la sélection de rues ou tronçons de rues, connues a priori pour leur densité commerciale, dans 6 zones-tests[14] (*Figure 19*) :

- **Metz-centre**, dans trois quartiers du centre-ville de Metz : il s'agit de secteurs commerçants de l'hypercentre historique, du quartier d'Outre-Seille ainsi que de rues du quartier du Sablon situé à proximité de la gare.
- **Metz-Actisud**, zone commerciale récente (créée en 2014) de la périphérie sud de Metz, située à cheval sur les communes d'Augny, Jouy-aux-Arches et Moulins (toutes trois communes de banlieue de l'Unité urbaine de Metz d'après l'INSEE). L'intégralité de la zone commerciale a été enquêtée.
- Deux « binômes » franco-allemands de villes voisines de la frontière : **Forbach (FR)** et **Völklingen (DE)**, situées à une dizaine de kilomètres à l'Ouest de Sarrebrück, villes moyennes de respectivement 20 000 et 40 000 habitants (données du LAU2, 2011). Et à une vingtaine de kilomètres à l'Ouest, **Creutzwald (FR)** et **Überherrn (DE)**, deux petites villes comptant respectivement 13 000 et 12 000 habitants dans les limites du LAU2 correspondant (2011).

Figure 19 – Les 6 zones d'enquête sélectionnées

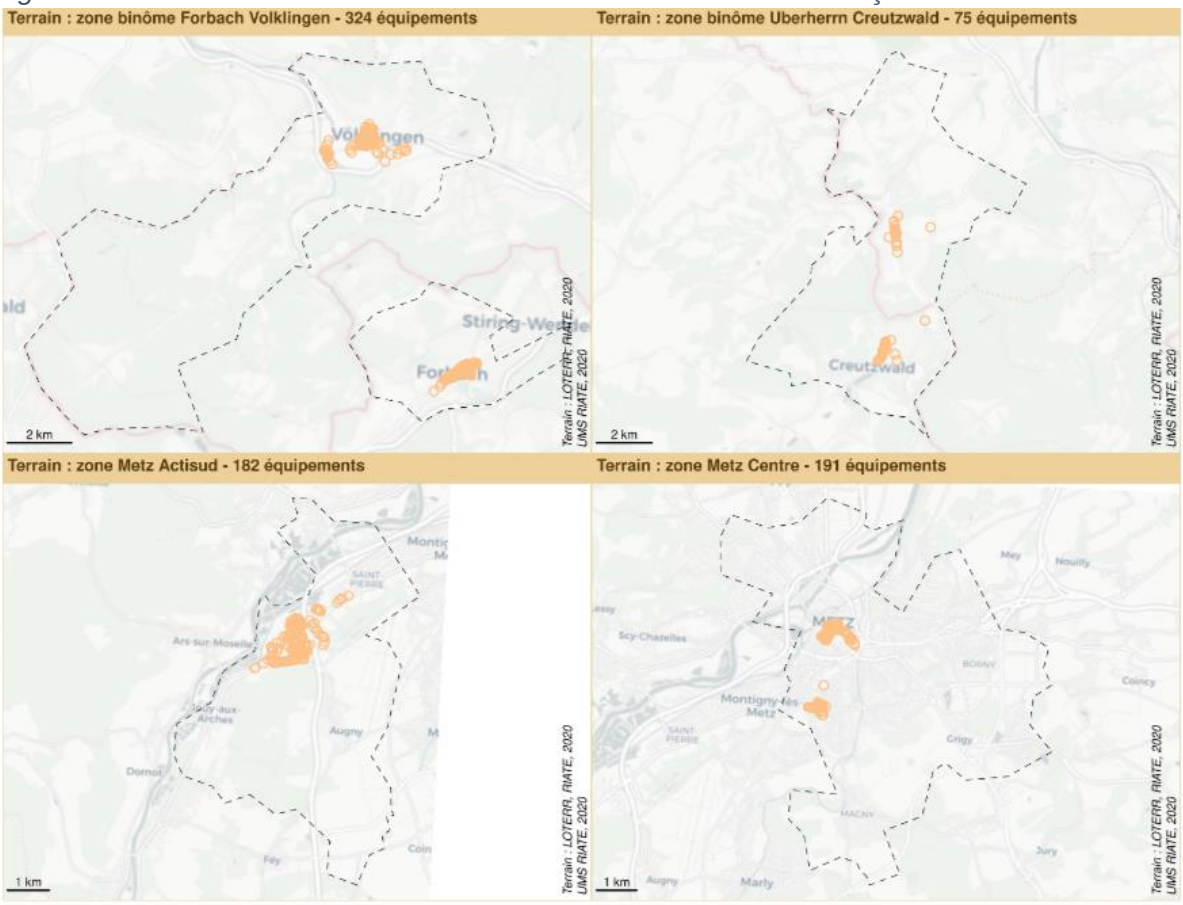


Dans le détail, les lieux d'enquête ne s'étendent pas sur l'ensemble du territoire communal mais ciblent une sélection de rues, de tronçons de rues, de zones commerciales, qui ont été

[14] Initialement 10 sites avaient été retenus, mais 4 zones qui devaient être traitées pendant le confinement ont dû être écartées de l'étude : il s'agissait des centres de Thionville au nord du Sillon Lorrain, de Sarrebrück en Allemagne, de Luxembourg-ville et de la zone commerciale d'Esch-sur-Alzette au Luxembourg.

identifiés par le LOTERR comme des espaces pertinents pour représenter différents types de forte concentration commerciale (*Figure 21*). Dans ces espaces, le protocole d'enquête s'appuie sur des relevés photographiques géolocalisés des 15 types d'équipements étudiés, qui sont intégrés à un système d'information géographique[15]. Dans la base de données associée, chaque équipement est décrit par deux types principaux d'attributs : des attributs de localisation (coordonnées X,Y) et des attributs de classification (selon la nomenclature du CNFG[16]). Les enquêteurs pouvaient ajouter des remarques à la base (par exemple sur la vacance d'un local commercial), à titre facultatif.

Figure 20 – Relevés de terrain dans trois villes françaises et deux villes allemandes



Intégration des relevés terrain et des données OSM dans une base commune

Pour chacune des zones d'enquête, l'intégration de données issues de différentes sources (relevés de terrain, OSM et la BPE dans le cas français) au sein d'une base commune est plus complexe qu'il n'y paraît. En effet, pour que la comparaison des équipements saisis dans les différentes sources ait un sens à cette échelle très fine d'observation, il est important de ne sélectionner que les équipements OSM et BPE correspondant aux rues ou aux zones précisément ciblées par l'enquête de terrain. Or cette **mise en correspondance spatiale** entre les équipements présents dans les deux (ou trois) bases peut être brouillée par la marge d'incertitude qui caractérise la localisation des équipements : sur le terrain, c'est à partir de la localisation de l'accès principal du bâtiment que les coordonnées sont saisies. Dans OSM, la

[15] Les relevés ont été directement saisis sur tablette informatique, à partir du logiciel QGIS.

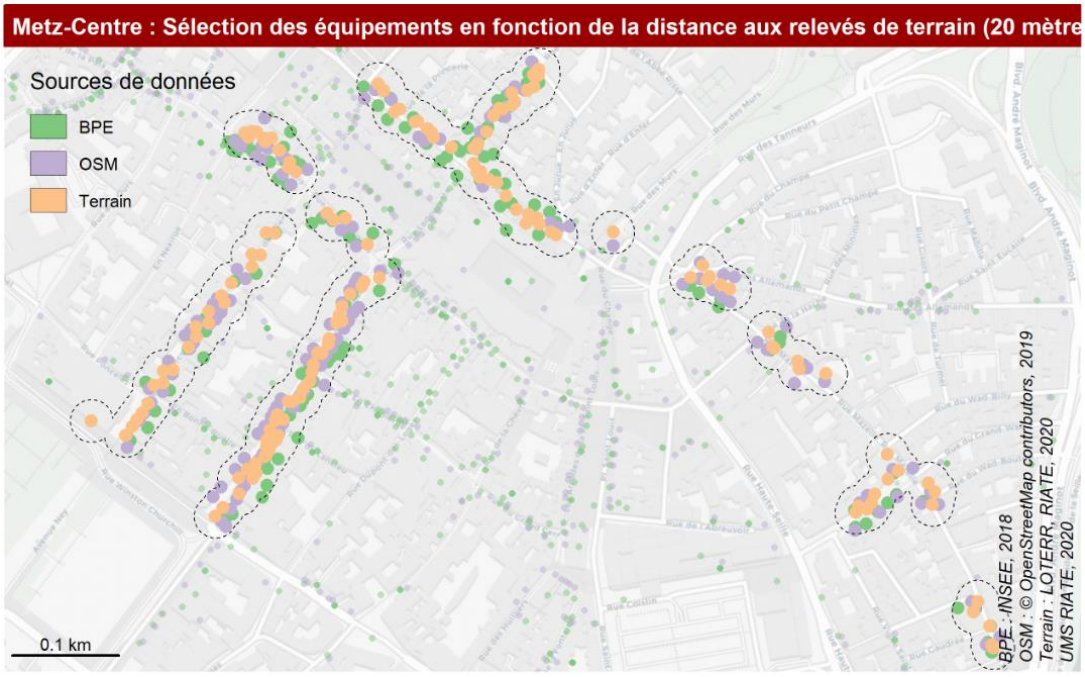
[16] Dans l'enquête de terrain, les attributs de classification des équipements s'appuient sur la nomenclature des commerces mise en place par la Commission commerce du Comité National Français de Géographie (CNFG), à partir de laquelle le LOTERR travaille pour ses relevés commerciaux.

localisation relevée par les contributeurs suit *a priori* une même logique, centrée sur les pratiques des utilisateurs, mais certains prétraitements peuvent modifier la précision de cette saisie, notamment lorsqu'il existe plusieurs accès (on fait alors la moyenne des coordonnées pour résumer la localisation à l'aide de coordonnées uniques). Dans la BPE, la localisation des équipements est connue à l'adresse avec une bonne précision dans l'absolu (écart avec la réalité du terrain inférieur à 100 m) mais cette marge d'incertitude n'est pas négligeable quand on travaille à l'échelle de la rue. De plus, l'affectation des coordonnées dans la BPE ne repose pas sur une méthode de saisie directe *in situ*, mais sur le fichier d'adresses du répertoire d'immeubles localisés (RIL).

Une fois la correspondance spatiale assurée, il a fallu **reconstituer une correspondance thématique entre les classifications OSM, BPE et CNFG,** afin de dénombrer les effectifs par catégories d'équipements de manière comparable d'une source à l'autre. Les catégories consolidées issues de la table de correspondance OSM/BPE (section 1.3) ont été associées à une nouvelle catégorie établissant la correspondance avec la nomenclature établie par le CNFG et utilisée par le LOTERR pour ses relevés de terrain.

Pour tenir compte de la marge d'incertitude résultant de l'hétérogénéité des sources, un espace tampon a été créé autour de chaque relevé de terrain, avec une portée de 20 mètres, retenue après plusieurs essais de seuils[17]. Cette portée de 20 mètres autorise *a priori* une capture suffisamment large pour englober les données OSM et BPE correspondant aux relevés terrains, sans pour autant « déborder » sur les autres équipements présents alentour. L'union de ces espaces tampons définit la zone d'extraction des équipements OSM (et BPE) à comparer aux relevés de terrain (*Figure 21*).

Figure 21 – Relevés de terrain, création d'une zone tampon et extraction des 15 catégories d'équipements dans OSM et dans la BPE : exemple dans l'hypercentre de Metz



[17] Dans les zones d'activités commerciales, où les emprises spatiales des équipements sont vastes et les écarts de localisation sont potentiellement plus importants d'une source à l'autre, cette portée de 20 mètres n'est pas suffisante. Une procédure spécifique a donc été retenue pour la zone commerciale d'Actisud : l'ensemble de la zone ayant été couverte par l'enquête, c'est le périmètre de la zone commerciale qui a servi de périmètre d'extraction des données OSM et BPE.

L’intégration thématique et spatiale des trois bases a permis de dresser des cartes d’inventaire d’équipements à la rue et d’évaluer les écarts entre les sources à un fin niveau de granularité spatiale.

Types d’erreurs récurrentes entre les sources

Une analyse détaillée de la position des points par source de données et catégories suggère l’existence d’erreurs récurrentes entre les sources, qui vont au-delà de la simple absence de saisie par les contributeurs OSM. Plusieurs situations types sont ressorties de l’observation d’extraits de rues présentant une hétérogénéité de situations et ont été validées par les enquêteurs sur le terrain :

- Situation 1 (*Figure 22*) : des problèmes liés à la méthode de mise en correspondance spatiale des équipements (zone tampon captant sur ses marges des équipements situés en dehors du périmètre d’enquête).
- Situation 2 (*Figure 23*) : des problèmes liés à une description ambigüe des catégories d’équipements entre OSM et la BPE.
- Situation 3 (*Figure 24*) : des problèmes de décalages temporels entre les versions de la BPE et d’OSM et/ou de mise à jour de la BPE (équipements anciens, toujours référencés dans la BPE et pas dans OSM).

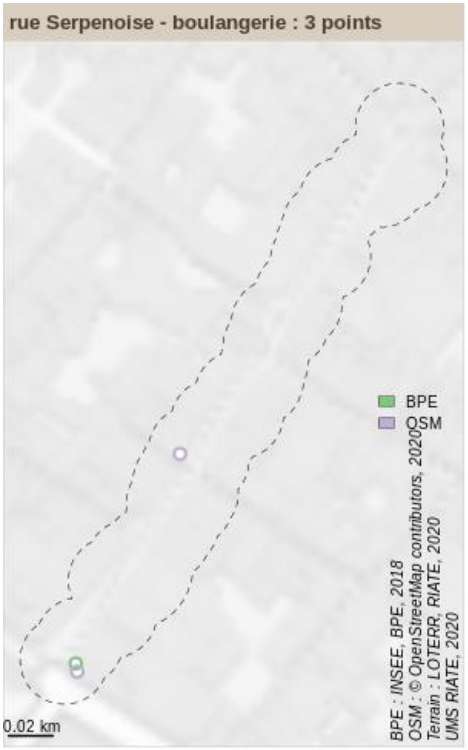
Figure 22 –Problèmes liés à la méthode de la zone tampon



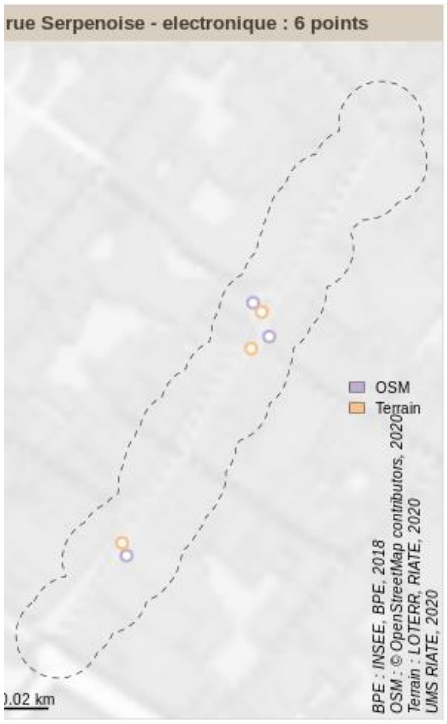
La pharmacie au nord de la rue Serpenoise appartient au centre commercial St Jacques qui n’a pas été inclus à la zone d’enquête



Figure 23 –Problèmes de définition de l'appartenance aux catégories



Une enseigne « Brioche dorée » qui n'est pas considérée comme une boulangerie dans la BPE ?

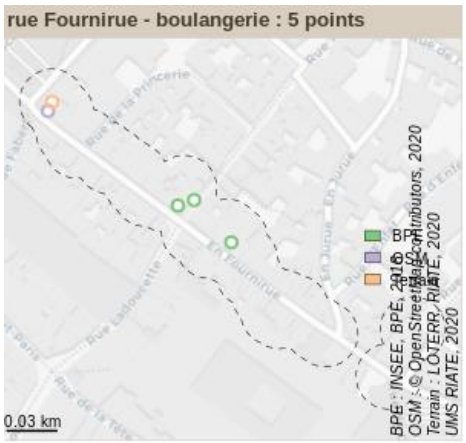


Des magasins de téléphonie mobile qui vendent du matériel électronique. Absents de la sélection de la BPE qui recense peut-être leur activité principale (opérateur téléphonique).

Figure 24 –Problèmes de décalages temporels OSM/BPE ou de mise à jour de la BPE



Le magasin de chaussures observé sur le terrain a ouvert très récemment. Il n'est pour l'heure pas référencé dans la BPE ou OSM. Le magasin au sud (BPE) n'existe plus. Le magasin au nord propose des chaussures, mais ce n'est pas sa vocation première (habillement).



La boulangerie au nord existe depuis très longtemps. Elle est non référencée dans la BPE. Au centre il s'agit d'un salon de thé qui propose des pâtisseries. Au sud il s'agit d'un établissement qui a fermé, remplacé aujourd'hui par un établissement de restauration.

3.2. Résultats : écarts globaux et par catégorie d'équipement

Analyse globale : le nombre d'équipements selon les sources

Pour l'ensemble des rues des 6 zones-tests et pour les 15 catégories étudiées, 382 équipements ont été dénombrés dans OSM, ce qui représente près de 50% des équipements relevés sur le terrain (772 relevés). On trouve un taux de complétude plus élevé que lorsqu'OSM était comparé à la BPE dans l'ensemble des communes françaises (le taux de complétude était de 31% dans ce cas), ce qui peut s'expliquer par la surreprésentation des zones de centre-ville. Mais cet écart varie fortement selon les sites (*Tableau 5*).

Tableau 5 - Le nombre d'équipements (15 catégories) dans les 6 zones tests

Nom LAU2	Pays	Terrain	OSM	BPE	Ecart Terrain/OSM (en %)**	Ecart Terrain/BPE (en %)
Metz	FR	191	164	217	**86**	114
Actisud*	FR	182	62	150	**34**	82
Forbach	FR	178	69	165	**39**	93
Völklingen	DE	146	81	NA	**55**	NA
Creutzwald	FR	46	1	27	**2**	59
Überherrn	DE	29	5	NA	**17**	NA
Total		*772*	*382*	*559*	***49,5***	*72,4*

** LAU2 correspondants : Augny, Jouy-aux-Arches et Moulins*
*** L'écart est mesuré de manière relative : (nbOSM/nbTerrain) * 100. Exemple de lecture : dans les lieux enquêtés à Metz centre, le nombre d'équipements OSM représente 86% du nombre de données terrain.*

Du côté français, la zone renseignée de la manière la plus complète dans OSM est celle du **centre de Metz**, où le nombre d'équipements présents dans OSM représente **86% des données terrain**, avec toutefois des variations sensibles selon les quartiers (de 78% aux Sablons à 103% dans l'hypercentre). Dans la zone commerciale récente de la banlieue sud de Metz (**Actisud**), la complétude d'OSM est plus faible et n'atteint que **34% des relevés de terrain**. Ce niveau est un peu inférieur à celui qui est observé dans le centre de **Forbach (39%)**. Enfin, il n'y a quasiment aucun relevé OSM recensé sur les lieux d'enquête de la petite ville de **Creutzwald (complétude de 2%**, 1 équipement OSM renseigné contre 47 sur le terrain). **Dans les deux zones d'enquête allemandes**, les données OSM sont nettement plus complètes que dans les villes françaises voisines : la **complétude est de 55% à Volklingen** (contre 39% à Forbach) et de **17% à Überherrn** (contre 2% à Creutzwald). Ce constat, bien que limité à deux localités, va dans le sens des analyses ayant souligné l'existence d'une communauté OSM plus ancienne et plus active en Allemagne.

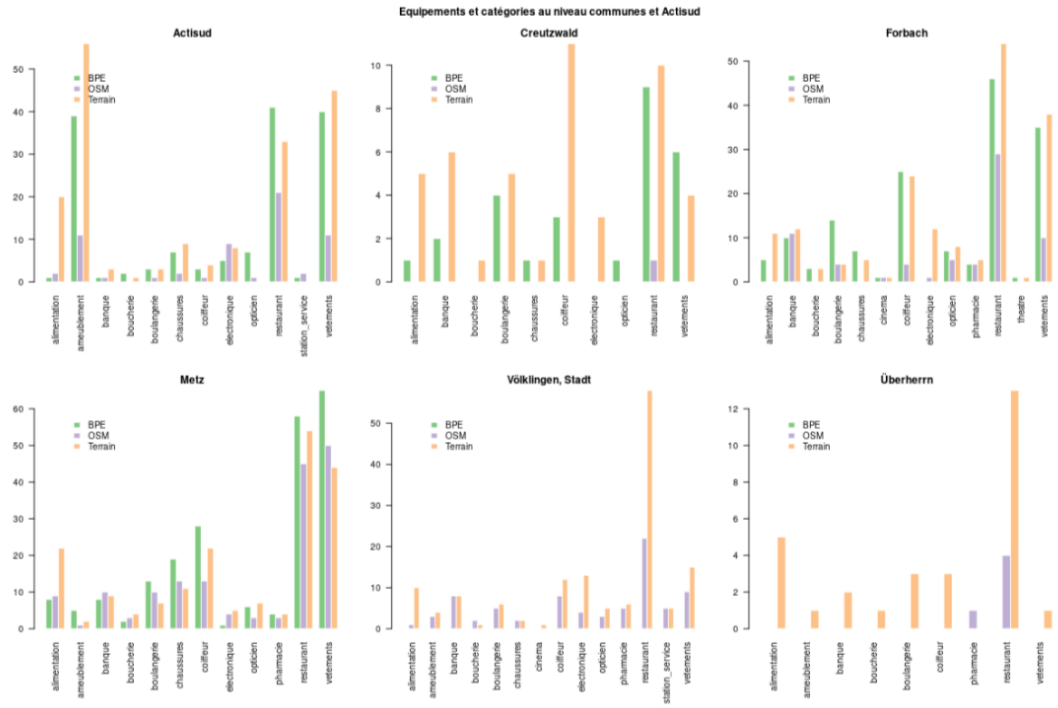
Enfin, en ce qui concerne la **BPE dans les zones-tests françaises**, on constate **parfois une surreprésentation des équipements BPE par rapport au terrain**, qui peut s'expliquer par le problème des "faux-actifs" déjà évoqué (équipements fermés qui resteraient enregistrés dans la BPE) : c'est le cas dans le centre de Metz, où l'on trouve en moyenne 14% d'équipements BPE en plus par rapport au terrain (entre 30 et 70% en plus dans certaines rues comme celles du Palais, des Clercs et du Petit Paris). Toutefois, **ce constat est loin d'être systématique** : dans les autres zones-tests, on constate plutôt une légère sous-représentation d'équipements dans la BPE (autour de 80% des relevés terrain, qui peut être due à des décalages observés dans les dates de relevés), parfois étonnamment plus marquée dans certaines localités (à Creutzwald les équipements BPE ne représentent que 59% des équipements de l'enquête de terrain - dans le détail seuls les cafés-restaurants, boulangeries, vêtements sont bien recensés).

Analyse par catégories d'équipements

Pour les zones d'enquête en France, on retrouve dans OSM les **mêmes écarts entre catégories plus ou moins bien représentées que ceux qui ressortaient de la comparaison OSM-BPE à l'échelle régionale** : au-delà des différences globales de complétude, qui dépendent des contextes spatiaux enquêtés, certaines catégories semblent relativement très peu présentes dans OSM (alimentation, ameublement, chaussures, coiffeurs notamment) tandis que d'autres se rapprochent davantage des effectifs constatés sur le terrain (banques, boulangeries, cafés-restaurants, magasins d'électronique, pharmacies…). Cela dit, le **profil du centre de Metz est assez spécifique**, au sens où certaines catégories habituellement sous-représentées comptent des effectifs plus nombreux que ceux des relevés terrain (par exemple les magasins de chaussures et de vêtements, les boulangeries), tandis que d'autres sont systématiquement sous-représentés, dans OSM comme dans la BPE (l'alimentation, par exemple).

L'**enquête à Völklingen** est particulièrement instructive du fait de l'absence de base de référence[18]. On peut relever un certain nombre de points communs avec les sites français : la complétude est ainsi faible à très faible pour les magasins d'alimentation (1 seul magasin recensé contre 10 sur le terrain) et relativement bonne à très bonne pour les banques (8 sur 8) et les pharmacies (5 sur 6). Mais les cafés-restaurants ne ressortent pas aussi bien que dans les sites français (24 sur 58) et certaines catégories habituellement sous-représentées côté français sont très complètes, à l'image des boulangeries (5 sur 6), des magasins d'ameublement (3 sur 4), de chaussures et de vêtement.

Figure 25 – Dénombrement des équipements recensés par terrain d'enquête, source de données et équipement



[18] La complétude pour Creutzwald est trop faible (2%) pour que les écarts entre catégories d'équipements puissent être interprétés. On pourra relever cependant que sur ce site, les seuls équipements recensés dans OSM sont des cafés-restaurants (et secondairement une pharmacie).

Conclusion partie 3

Dans les 6 zones enquêtées, l'intégration thématique et spatiale des bases OSM, des relevés de terrain et de la BPE (pour les sites français) a permis de dresser des cartes d'inventaire d'équipements à la rue et d'évaluer les écarts entre les sources à un niveau fin de granularité spatiale. Parmi les sites d'enquête en France, si la zone renseignée de la manière la plus complète dans OSM est sans surprise celle du centre de Metz, la zone commerciale récente de la banlieue sud de Metz ne bénéficie pas d'une meilleure couverture dans OSM que la moyenne (1/3 des équipements réels recensés dans OSM). D'autre part, la comparaison entre la BPE et les relevés de terrain a fait parfois ressortir (notamment dans le centre de Metz) une surreprésentation des équipements BPE par rapport au terrain, qui peut s'expliquer par le problème des "faux-actifs" déjà évoqué plus tôt ; mais ce constat est loin d'être systématique et ne peut être généralisé à l'ensemble des sites étudiés.

Dans les deux zones d'enquête allemandes, les données OSM sont nettement plus complètes que dans les villes françaises situées de l'autre côté de la frontière. Ce constat, bien que limité à deux localités, va dans le sens des analyses ayant souligné l'existence d'une communauté OSM plus ancienne et plus active en Allemagne. Le cas mieux documenté de Völklingen a permis de dresser un premier aperçu des équipements les plus contribués dans OSM du côté allemand, qui recoupe partiellement les conclusions des sites d'enquête français : les catégories bien contribuées simultanément de part et d'autre de la frontière demeurent les banques et les pharmacies, mais les cafés-restaurants ressortant sensiblement moins que du côté français. Ces premières pistes de comparaison internationale devraient être complétées et consolidées en élargissant la sélection à d'autres zones d'enquêtes allemandes et luxembourgeoises, caractéristiques d'autres contextes spatiaux (notamment les centres-villes ou zones commerciales de plus grandes agglomérations).

4. Identifier des pôles de concentration d'équipements OSM à l'échelle transfrontalière

Les analyses précédentes ont permis d'identifier les équipements présentant la plus grande fiabilité dans la base OSM, parmi les 15 catégories retenues, à la fois du fait de la clarté de leur définition (pas ou peu d'ambiguïté dans la classification d'OSM) et de leur bonne complétude par rapport à la BPE et aux relevés de terrain. Partant d'une sélection de 4 équipements de proximité (cafés-restaurants, banques, pharmacies) et intermédiaires (magasins d'électronique)[19], il s'agit à présent d'analyser leur répartition spatiale à l'échelle transfrontalière, afin d'**identifier des zones de forte concentration relative d'équipements** et de **voir comment la hiérarchie de pôles qui s'en dégage varie selon les choix méthodologiques effectués**. Précisons d'emblée que cette partie a une portée avant tout exploratoire et vise à tester une méthode d'identification de pôles d'équipements à partir de relevés ponctuels, dans l'hypothèse qu'OSM soit renseignée de façon harmonisée. Cette harmonisation (internationale, notamment) n'étant pas encore vérifiée, l'interprétation thématique des résultats doit être menée pour le moment avec une grande prudence.

Nous présentons tout d'abord les choix méthodologiques permettant une lecture comparable des concentrations locales d'équipements de part et d'autre des frontières : l'utilisation d'unités statistiques locales de référence (LAU2), de tailles très hétérogènes dans l'espace transfrontalier, risque d'introduire d'importants biais cartographiques et statistiques, connus sous le nom de MAUP (« Problème des unités spatiales modifiables », Openshaw, 1981). Afin d'éviter ces biais, la démarche retenue s'appuie sur un dénombrement des équipements au sein d'un carroyage régulier (grille kilométrique) et sur la mesure d'un potentiel local de ressources dans le voisinage de chaque lieu (méthode des potentiels de Stewart). Afin d'analyser les inégalités de répartition des équipements dans l'ensemble de l'espace d'étude, nous testons ensuite l'hypothèse selon laquelle ce potentiel dépend de la demande locale, en mesurant la relation statistique entre le potentiel d'équipements et le potentiel de population. Nous prêterons une attention particulière à la répartition géographique des écarts à cette relation, afin de discuter de la sous ou sur-représentation d'équipements, toutes choses égales quant à la population. Enfin, nous proposons une méthode d'identification de centralités commerciales locales qui repose sur l'identification de pôles de concentration relative d'équipements par rapport à leur environnement proche.

4.1. Mesurer des potentiels d'équipements dans le voisinage des lieux

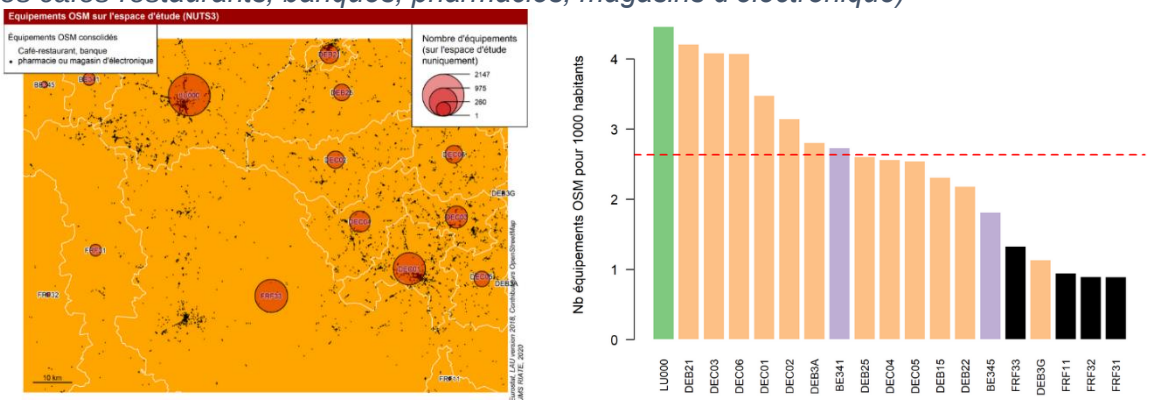
Densités d'équipements au niveau des NUTS3 et des LAU2

L'espace d'étude transfrontalier compte plus de 8300 points d'intérêts OSM correspondant aux 4 équipements sélectionnés[20]. Une première vision très globale de leur répartition spatiale peut être présentée au niveau des NUTS3[21], qui correspond au niveau des départements français (*Figure 26 – Densité des équipements OSM pour 1000 habitants, à l'échelon des NUTS3 (pour les cafés-restaurants, banques, pharmacies, magasins d'électronique)Figure*

[19] Les théâtres et cinémas, équipements de rayonnement régional, n'ont pas été conservés dans cette partie.
[20] Dans le détail des catégories, 78% de ces POI correspondent aux cafés-restaurants, puis 11% aux banques.
[21] Ou des parties de NUTS3, pour les mailles qui ne sont pas complètement incluses dans la zone d'étude.

26). Les densités d'équipements, exprimées en nombre d'équipements pour 1000 habitants, opposent ainsi les zones les plus équipées, situés au Luxembourg et en Allemagne – d'abord dans les NUTS3 de Trèves (DEB21) au nord, de St-Wendel (DEC06) et Neunkirchen (DEC03) à l'est de Sarrebrück, puis de Sarrebrück même (DEC01) et de Merzig-Wadern -, avec plus de 4 équipements pour 1000 habitants, à l'ensemble des départements français de l'espace d'étude (ainsi que le NUTS tronqué de Kusel - DEB3G), qui comptent chacun environ 1 équipement pour 1000 habitants.

Figure 26 – Densité des équipements OSM pour 1000 habitants, à l'échelon des NUTS3 (pour les cafés-restaurants, banques, pharmacies, magasins d'électronique)

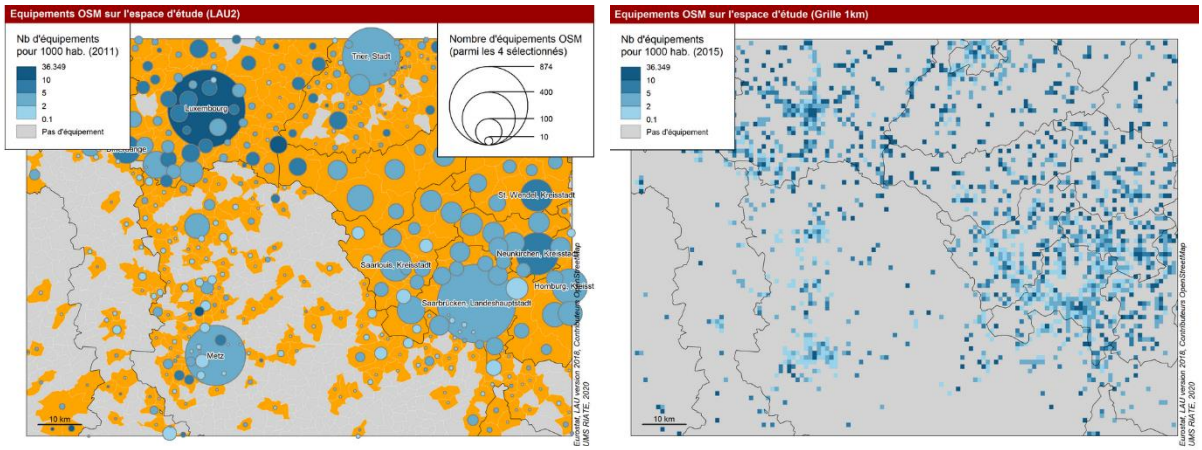


a. Présence d'équipements OSM et nombre d'équipements par NUTS3

b. Densité d'équipements OSM, pour 1000 habitants, par NUTS3

A l'échelon plus fin des municipalités (LAU2, *Figure 27.a*), la densité d'équipements fait ressortir une répartition spatiale de nature beaucoup plus lacunaire et fragmentée côté français, mais qui reflète avant tout les fortes inégalités de taille des mailles statistiques locales, beaucoup plus petites en France qu'au Luxembourg et en Allemagne. Pour s'abstraire de ce maillage statistique très hétérogène de part et d'autre des frontières, le découpage spatial privilégié est un carroyage régulier, calé sur la grille kilométrique d'Eurostat, pour laquelle on dispose aussi de la donnée de population totale (en 2011). Il est ainsi possible de rapporter le nombre d'équipements au nombre d'habitants au sein d'unités spatiales homogènes (*Figure 27.b*). Toutefois, cette représentation des densités d'équipements au carreau suppose que les équipements en question ne desservent que la population résidant dans les limites de cette zone très restreinte, sans prendre en compte les interactions avec l'environnement proche.

Figure 27 – Densité d'équipements à l'échelon local (pour les cafés-restaurants, banques, pharmacies, magasins d'électronique)



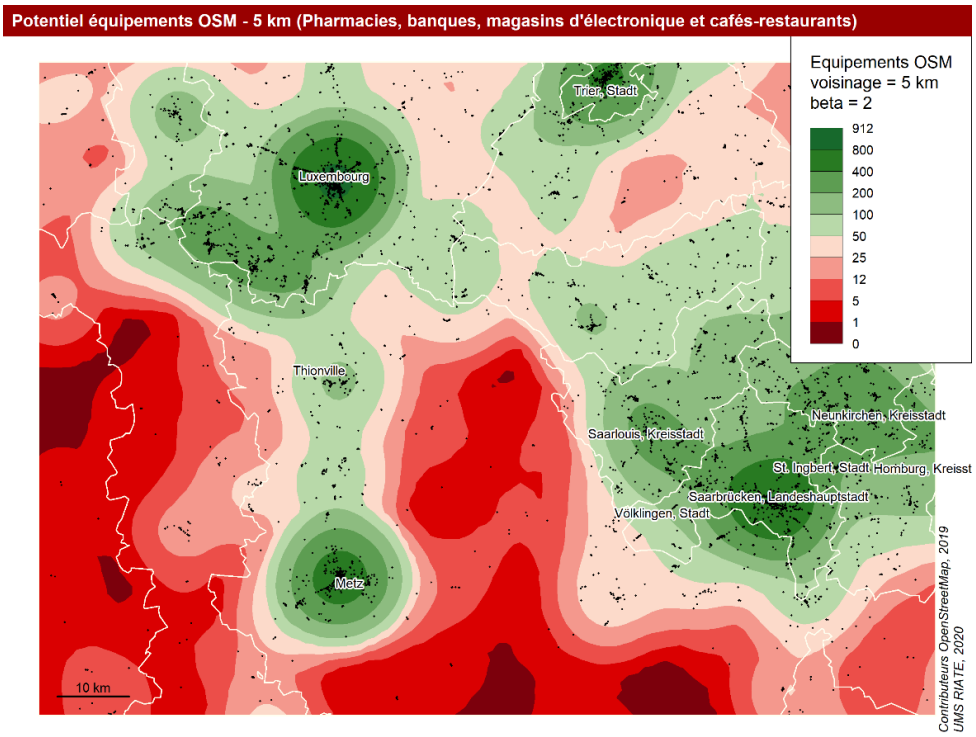
a. au niveau des LAU2

b. dans un carroyage de 1 km

Potentiels d'équipements dans un voisinage proche

La représentation de la présence d'équipements dans le voisinage de chaque lieu repose sur la méthode de calcul de potentiel de Stewart (Stewart, 1942). Cette méthode consiste à évaluer les quantités d'équipements disponibles non seulement au sein de chaque lieu (ici chaque carreau de la grille) mais aussi autour de ce lieu, jusqu'à une certaine portée géographique. Pour chaque carreau de l'espace d'étude, le nombre d'équipements présents dans le voisinage (par exemple à 5 km ou 10 km autour de chaque carreau) est pondéré par la distance aux carreaux équipés, selon une fonction décroissante de la distance qui exprime la diminution du potentiel de déplacement vers les équipements avec l'éloignement. La *Figure 28***Erreur ! Source du renvoi introuvable.** représente par exemple le potentiel d'équipements OSM dans un voisinage de 5 km, la

Figure 28 – Potentiel d'équipements OSM dans un voisinage de 5 km (cafés-restaurants, banques, pharmacies, magasins d'électronique



décroissance des probabilités d'interaction avec la distance étant estimée à l'aide d'une fonction exponentielle négative. Les zones où le potentiel d'équipement est le plus élevé (plus de 800 équipements dans une portée de 5 km) sont plus ou moins étendues et font ressortir les pôles de Luxembourg, de Saarbrück, puis ceux de Metz et de Trèves. On distingue aussi des régions bien équipées (de 400 à 800 équipements) au sud du Luxembourg de part et d'autre d'Esch-sur-Alzette ainsi que dans le sud de la Sarre, entre Saarlouis et Neunkirchen. De part et d'autre du sillon lorrain, à l'inverse, la situation dominante est celle d'une très faible présence voire d'une absence d'équipements dans un voisinage proche.

4.2. Analyse de la relation entre les potentiels locaux d'équipements et de population

Ces potentiels d'équipements mesurés d'après la base OSM ne pouvant pas être comparés à des bases d'équipement de référence dans les parties allemande, luxembourgeoise et belge de l'espace d'étude, nous les comparons à la population qui tient ici lieu de variable « de contrôle » indirecte. On cherche à tester l'hypothèse selon laquelle le potentiel d'équipements dépend en grande partie de la demande locale (potentiel de population) dans un voisinage donné, à l'aide d'un modèle statistique de régression linéaire. Il s'agit donc d'analyser l'intensité de la relation entre concentrations d'équipements et de population et d'accorder une attention particulière aux écarts à cette relation. S'il existe une relation significative, en effet, l'analyse de la distribution spatiale des résidus de la régression devrait permettre de mettre en évidence les espaces où les équipements OSM sont sous-représentés / sous-contribués, toutes choses égales quant à la population. On notera d'emblée que l'interprétation des écarts à cette relation peut être tout à fait équivoque : si la complétude de la base OSM était parfaite, ces écarts permettraient d'identifier les tendances locales à la sur ou sous-représentation des équipements par rapport à la population résidente. De même, si la densité d'équipements par habitant était homogène dans l'ensemble de l'espace d'étude, la population pourrait servir de donnée de contrôle pour identifier les zones les mieux et les moins bien contribuées dans OSM. Faute de donnée supplémentaire, l'interprétation des résidus s'inscrit dans cette

ambiguïté et appelle à une certaine prudence. Les analyses seront parfois menées dans chacun des périmètres nationaux de la zone d'étude pour tenter de neutraliser les effets potentiels de différences nationales de contribution à OSM dans l'interprétation des écarts.

La relation entre le potentiel de population et le potentiel d'équipements[22], pour une portée locale de voisinage de 5 km, montre une qualité d'ajustement élevée (R^2 = 72%) et significative (

Figure 29). Le graphique correspondant souligne néanmoins la moindre intensité de cette relation pour les zones les moins peuplées (population inférieure à 10 000 habitants dans le voisinage).

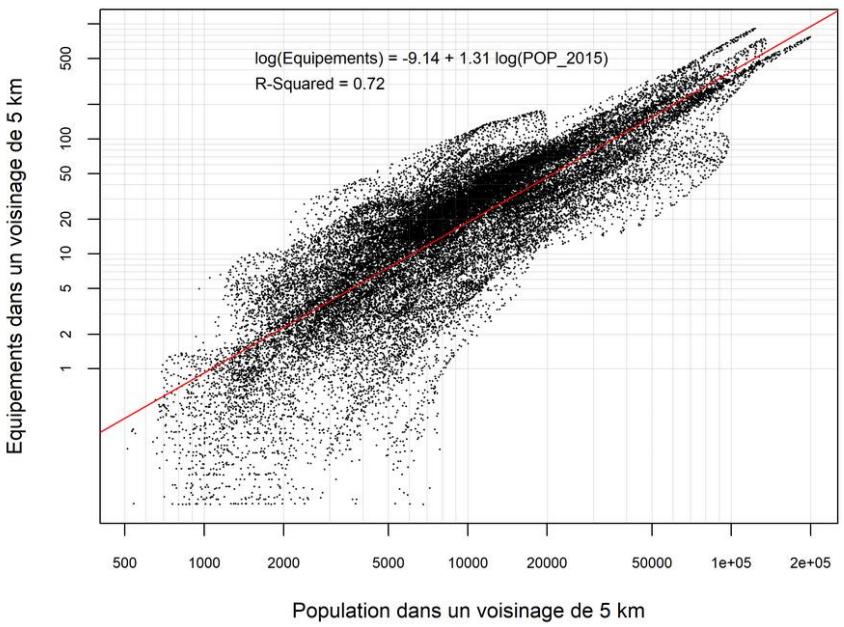


Figure 29 – Modèle de régression linéaire entre le nombre d'équipements et la population résidente dans un voisinage de 5km.

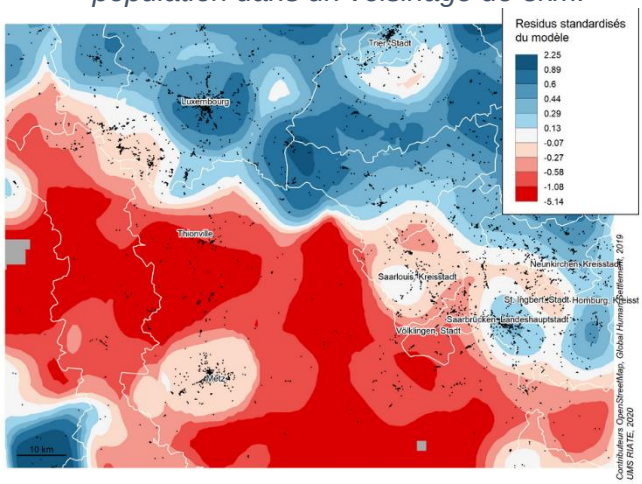
La cartographie des résidus à cette relation permet d'interroger l'organisation spatiale qui les sous-tend (*Figure 30.a*). Au niveau de l'ensemble de l'espace d'étude, elle révèle une forte discontinuité entre la France et les pays voisins. Mis à part l'espace environnant Metz et l'extrême sud-ouest de l'espace d'étude, on observe significativement moins d'équipements OSM dans la partie française de l'espace d'étude (en bleu, zones où les équipements sont moins nombreux que ne le laisserait supposer la population). Le même modèle testé dans des limites nationales (*Figure 30.b* à *Figure 30.d*) montre des variations sensibles des résultats de l'ajustement : sa qualité moins élevée en France qu'en Allemagne et au Luxembourg (R^2 respectivement de 75%, 91% et 94%) souligne la plus faible dépendance de la relation entre population et équipements et rappelle l'existence de zones urbaines peuplées montrant un déficit relatif d'équipements dans OSM (par exemple à Hayange et Sarreguemines). La valeur

[22] La relation testée est la suivante : log(nombre d'équipements) = a * log(population) + b

de la constante, plus faible en France, témoigne du moindre niveau d'équipement qui ressort d'OSM, tandis que le coefficient directeur de la droite plus élevé souligne une tendance à l'existence de contrastes d'équipements plus élevés entre les zones faiblement peuplées et les zones de forte concentration démographique.

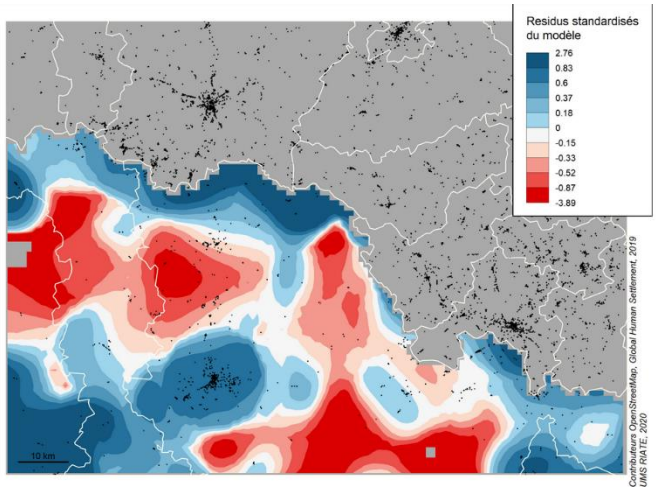
En raison d'une forte discontinuité entre la France et les Etats voisins, qui pourrait s'expliquer par une moindre contribution à OSM côté français et/ou par une densité d'équipements globalement plus faible par rapport à la population résidente, la représentation des potentiels d'équipement invite donc à la prudence et nous amène à privilégier une mesure relative des concentrations d'équipements, qui s'appuie sur les différentiels locaux de potentiels, plutôt qu'une mesure des niveaux absolus d'équipements présents dans le voisinage des lieux.

Figure 30 – Résidus des modèles de régression linéaire entre le potentiel d'équipements et de population dans un voisinage de 5km.



a. Ensemble de l'espace d'étude

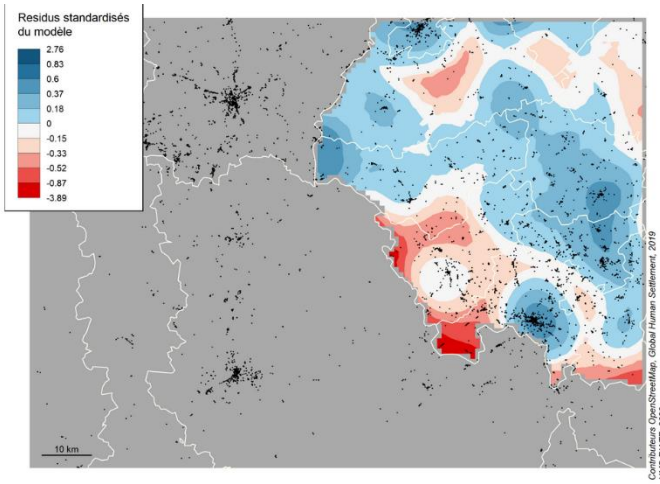
Log(Equip)=-9,1 + 1,3 * Log(Pop) (R^2=72%)



b. Partie française

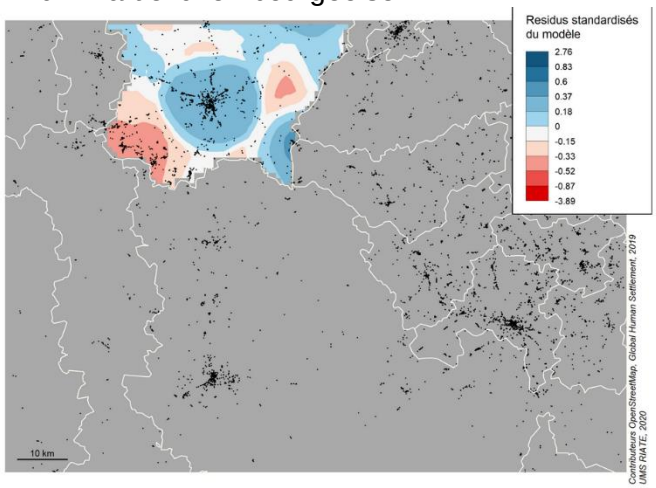
Log(Equip)=-10,22 + 1,3 * Log(Pop) (R^2=75%)

c. Partie allemande



Log(Equip)=-6 + 1,03 * Log(Pop) (R^2=91%)

d. Partie luxembourgeoise



Log(Equip)=-7,8 + 1,2 * Log(Pop) (R^2=94%)

4.3. Identifier des pôles de concentration relative d'équipements

Pics locaux de concentration d'équipements

L'identification de centralités de ressources (de commerces, de populations…) peut, parmi de nombreuses méthodes, reposer sur la comparaison des équipements disponibles au sein de voisinages de différentes portées, pour identifier des zones de concentration relative par rapport à leur environnement proche. Cette méthode a déjà été utilisée dans des travaux précédents afin de faire émerger des pôles transfrontaliers en fonction de la concentration spatiale relative de certaines ressources (De Ruffray, Hamez, 2008). La centralité est donc appréhendée ici à partir d'une définition mettant l'accent sur la concentration d'une offre de biens et de services qui polarise une population plus dispersée (Hypergeo, 2004). Un pôle de concentration relative de ressources est défini comme un lieu doté d'un potentiel d'équipements plus important que dans un voisinage plus large. Plus précisément, la mesure utilisée repose sur le rapport entre le potentiel d'équipements à courte distance et le potentiel à plus longue distance (pot_dist(min)/pot_dist(max)). Transposé à la question des équipements présents dans *OSM* et à la problématique générale des petites et moyennes centralités, la mise en relation de ces concentrations d'équipements à différentes portées spatiales peut faire émerger des situations particulières :

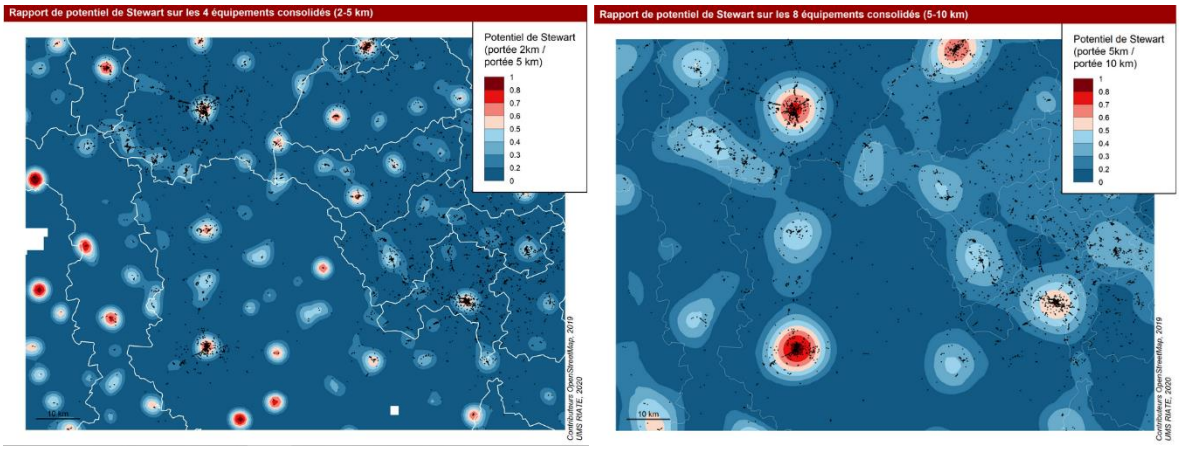
- de zones de concentration relative d'équipements plus forte que dans le voisinage immédiat de cette zone (ratio voisinage proche / voisinage plus éloigné élevé).

- de zones où les équipements sont moins nombreux que dans le voisinage (ratio voisinage proche / voisinage plus éloigne faible)[23].

Dans les cartes qui suivent, les pôles qui ressortent en tons chauds sont ceux qui correspondent à une valeur supérieure à 0.5 : la zone centrale (c'est-à-dire située dans un voisinage de courte portée autour de l'unité spatiale observée, par exemple 2 km pour la carte de gauche) concentre la moitié des équipements présents dans un voisinage élargi de 5 km (ou autrement dit, que pour 1000 équipements présents dans un voisinage de 2 km, et 2000 dans un voisinage de 5 km, le rapport vaudra 0,5).

Figure 31 – Pics relatifs de centralité : rapport entre les potentiels d'équipements mesurés entre 2 portées (pour les cafés-restaurants, banques, pharmacies, magasins d'électronique)

[23] Notons qu'avec cette méthode, une absence de pic de concentration d'équipements n'est pas forcément synonyme d'absence d'équipements. Un rapport proche de zéro signifie simplement qu'il n'existe pas de concentration géographique d'équipements à cet endroit (répartition homogène des équipements OSM). Cela peut concerner à la fois des espaces peu équipés de façon homogène spatialement (nord-est de la Moselle par exemple) ou à l'inverse, qui concentrent les équipements sur un espace relativement étendu. C'est par exemple le cas du Sillon Lorrain situé entre Thionville et Metz, qui voit un léger pic de concentration localisé sur Amnéville, mais qui se distingue peu de son voisinage géographique proche.



a. Entre des portées de 2 et 5 km

b. Entre des portées de 5 et 10 km

Le choix de la portée des lissages influence grandement les résultats : la carte des centralités construite à partir du rapport entre les potentiels de 5 km et de 10 km de portées fait ressortir les 4 principales aires urbaines de l'espace d'étude, tout en soulignant dans les espaces intermédiaires le contraste entre des formes de répartition de ces équipements de proximité plutôt concentrées en France et dispersées en Allemagne. Rappelons toutefois que la validité de ces résultats repose sur l'hypothèse selon laquelle les contributions OSM sont homogènes dans les voisinages géographiques sur lesquels portent ces calculs de potentiel. On supposera ici que c'est le cas pour la mise en pratique méthodologique. Mais nous émettons d'importantes réserves, notamment en raison de l'hétérogénéité locale des contributions OSM déjà soulignée, par exemple entre Forbach et Vörklingen, ou encore entre Metz-centre et sa zone commerciale Actisud, voire parfois la quasi absence de contribution, comme à Creutzwald qui présente pourtant des caractéristiques de centralité (au-delà de la rue commerçante du centre, présence d'un parc d'activité au sud de la commune, où sont simplement référencés en septembre 2020 dans OSM un centre commercial Leclerc et des services dédiés à l'automobile).

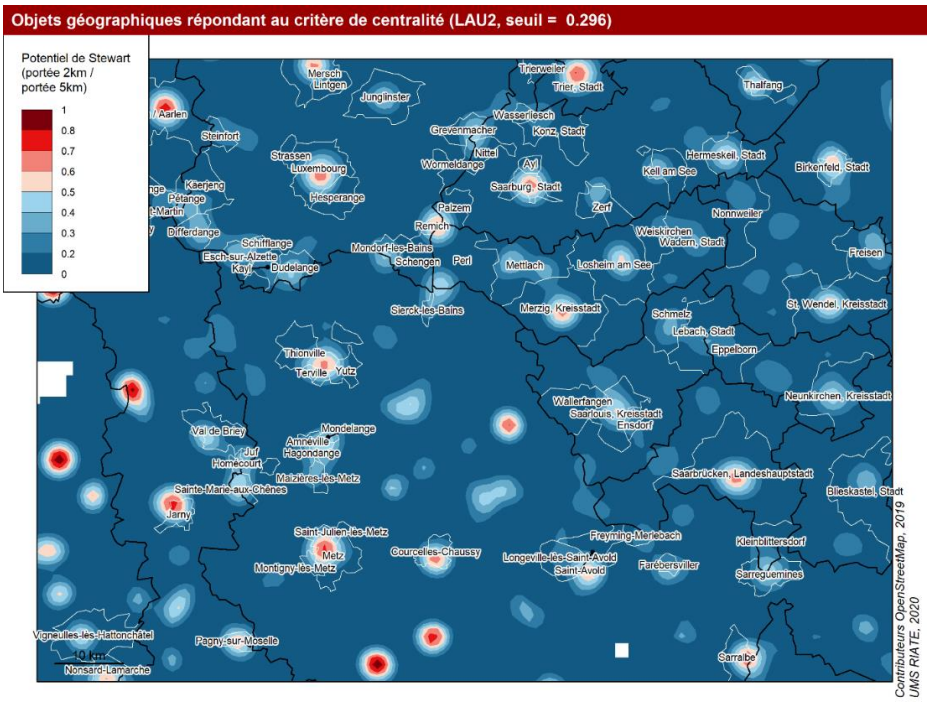
Typologie des municipalités en termes de centralités d'équipements

Ces pics de concentration d'équipements étant générés dans un espace continu à l'aide de représentations lissées, il est peu évident d'apprécier la réalité administrative (via l'affectation d'un nom, par exemple) à laquelle ils se rattachent. Une méthode a été proposée pour identifier et représenter les unités spatiales locales dont les limites administratives croisent ces pics de concentration.

Dans un premier temps, les communes qui croisent les petits pics de centralité sont identifiées, à partir d'une valeur seuil du rapport de potentiel de 0.3, qui fait ressortir 10 % des valeurs les plus élevées de l'espace d'étude (figure 4.6). Autrement dit, dans cet espace défini par ces deux voisinages, plus de 30 % des équipements d'un voisinage de 5 km sont concentrés dans un voisinage de 2 km. D'autre part, on considère qu'une commune doit disposer d'au minimum 5 équipements OSM pour être définie comme centrale. Le fait d'établir un seuil minimal à 5 équipements permet notamment d'exclure des espaces de très faible concentration absolue d'équipements, en Meuse et Meurthe-et-Moselle notamment. En revanche, ce paramètre permet de conserver parmi les pics locaux remarquables des zones pour lesquelles le rapport de potentiels est plus faible mais où la quantité d'équipements est globalement importante : c'est le cas par exemple des conurbations à la frontière luxembourgeoise et au niveau

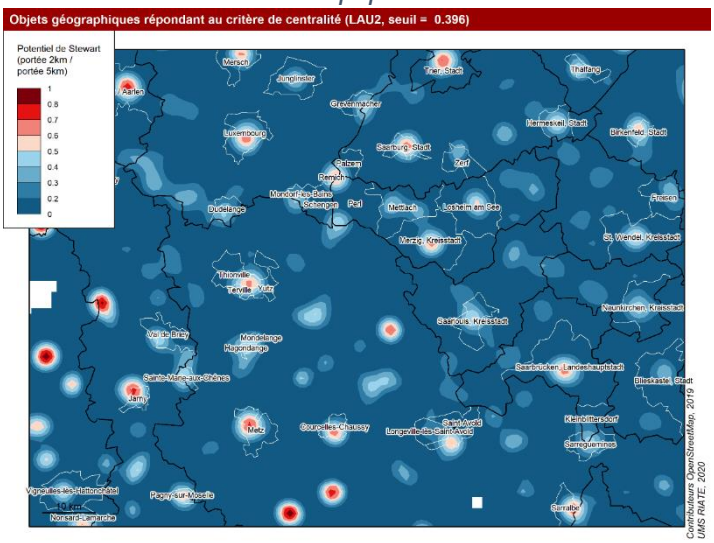
d'Amnéville, qui présentent une répartition d'équipements plutôt homogène dans un voisinage proche.

Figure 32 – Pics locaux de concentration d'équipements – 10 %



Pics nommés s'ils appartiennent aux 10 % des valeurs les plus élevées des concentrations d'équipements – communes contribuées par au moins 5 équipements OSM de type cafés-restaurants, banques, pharmacies ou magasins d'électronique. Les valeurs très prononcées à 30 km au sud et à l'ouest de Metz ne sont pas nommées car les communes correspondantes comptent moins de 5 équipements.

Figure 33 – Pics locaux de concentration d'équipements – 5 %

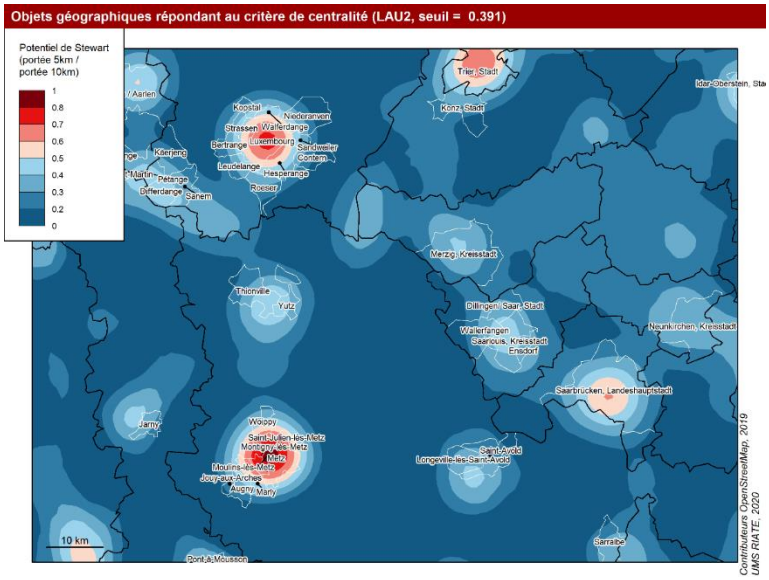


5 % des valeurs les plus élevées des concentrations d'équipements – communes contribuées par au moins 5 équipements OSM de type cafés-restaurants, banques, pharmacies ou magasins d'électronique.

En augmentant le seuil à 0.4 (rapport de voisinage correspondant au 5 % des valeurs observées sur l'espace d'étude) on visualise à présent les communes caractérisées par une concentration d'équipements dans un rapport de voisinage (2 km / 5 km) supérieur à 40 %.

Le fait d'élargir la portée de ce rapport de potentiel (5-10 km contre 2-5 km précédemment) produit une toute autre image. Ce contexte géographique d'analyse illustre ici davantage des conurbations présentant des pics de centralité. Cette portée de lissage est d'ailleurs sûrement plus appropriée pour comparer convenablement les LAU2 français et ceux des pays voisins et contourner les effets de MAUP évoqués plus haut. On s'intéresse ici aux 10 % des valeurs les plus élevées dans ce contexte géographique (seuil 0.39). Les communes qui croisent ces pics doivent comporter plus de 10 équipements OSM pour être considérées comme présentant des caractéristiques de centralité dans ce contexte géographique

Figure 34– Pics régionaux de concentration d'équipements – 10 %



10 % des valeurs les plus élevées des concentrations d'équipements – communes contribuées par au moins 10 équipements OSM de type cafés-restaurants, banques, pharmacies ou magasins d'électronique

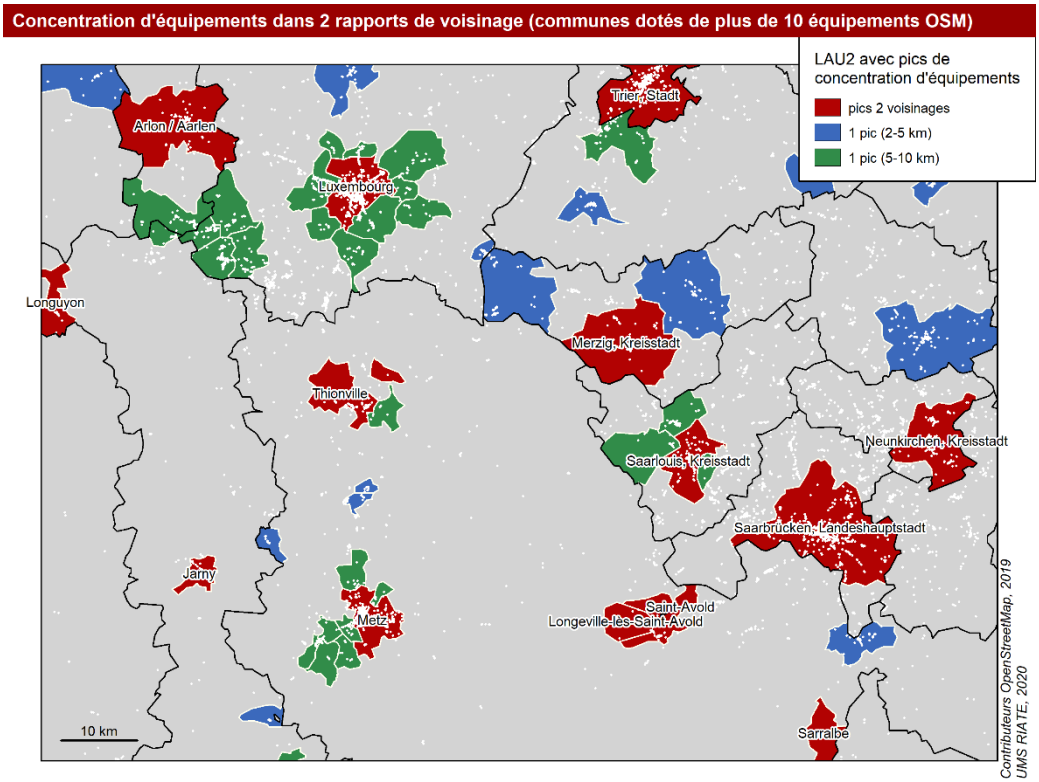
Après avoir discuté de l'importance de ces paramètres pour l'identification de potentielles centralités d'équipements frontalières, on propose ici une typologie de synthèse qui identifie les communes qui intersectent des pics de centralité (5 % des valeurs les plus élevées) pour les 2 groupes de paramètre suivants :

- **Centralité régionale** : Rapport 5/10 km, seuil 0.391, 10 équipements OSM au minimum (ensemble des équipements consolidés).
- **Centralité locale** : Rapport 2/5 km, seuil 0.458, 10 équipements OSM au minimum (ensemble des équipements consolidés).

Le croisement de ces deux indicateurs de centralité aboutit à une typologie où l'on identifie les communes qui croisent des pics de centralité au niveau régional et local (en rouge sur la carte), pour aucun des deux critères (en gris), pour le niveau de centralité régional uniquement (en vert), et pour le niveau de centralité local uniquement (en vert).

La surimposition des points OSM sur cette carte de synthèse souligne les apports et les limites de la méthode : si elle s'avère a priori robuste pour identifier les principales concentrations, elle semble en revanche largement améliorable pour qualifier les espaces densément équipés de façon relativement homogène dans l'espace. C'est notamment le cas à l'est de la conurbation franco-luxembourgeoise, où on peut observer visuellement une certaine concentration de points, que cette typologie (avec les seuils retenus) ne fait ici pas ressortir. Par ailleurs, l'effet de MAUP lié à l'hétérogénéité des mailles LAU2 (communes sarroises nettement plus grandes) est réintroduit via le croisement avec les municipalités, et appellerait sans doute la prise en compte d'autres maillages côté français (par exemple les EPCI).

Figure 34 – Pics régionaux et locaux de concentration d'équipements, tentative de croisement des approches



Conclusion partie 4

A partir d'un semis d'équipements incomplet, nous avons identifié des concentrations dans un espace continu proche géographiquement, puis proposé, avec une intersection simple, d'extraire ces concentrations dans espace discontinu (un maillage territorial), support de la prise de décision. Même si la hiérarchie des centralités ne peut être pour l'instant dégagée des données OSM, du fait de l'importance de leurs lacunes, cette méthode est prometteuse en ce qu'elle donne une approximation raisonnable des principales centralités commerciales qui structurent ces espaces d'étude : Sarrebrück, Metz, Luxembourg, Trèves et dans une moindre mesure Thionville ressortent assez nettement des analyses, avec pour seule source de données la base OSM, pour les équipements de proximité les plus fiables de la base. On peut se demander aussi dans quelle mesure on ne produit pas ici en définitive – et en l'état actuel du projet OSM-, une cartographie des concentrations non pas de quelques équipements commerciaux (au premier rang desquels les cafés-restaurants) mais des contributions OSM, ce qui pourrait s'assimiler à une géographie des centralités du réseau des utilisateurs d'OSM. Cette analyse exploratoire doit donc être comprise comme une première pierre à l'édifice pour envisager les centralités commerciales grâce à OSM. Le fait de disposer de scripts d'analyse accessibles et reproductibles ouvre les résultats à la vérification et à la discussion.

Conclusion

Cette étude a permis de documenter la complétude de l'information sur les équipements dans OSM, appliquée ici au cas de la Grande région, mais transposable à d'autres zones géographiques. Il en ressort que l'information OSM sur les équipements est en général plutôt fiable dans les centres-villes densément peuplés, plus aléatoire dans les périphéries urbaines et les zones rurales hors influence urbaine. Cette hétérogénéité spatiale à l'échelle régionale se double d'une importante hétérogénéité entre les pays de la zone d'étude, avec une forte discontinuité reflétant le niveau plus faible des contributions du côté français, par rapport aux pays voisins. En l'état actuel, la base OSM ne peut donc compenser l'absence de bases nationales de référence sur les équipements en Allemagne, au Luxembourg et en Belgique, sauf pour les zones centrales des plus grandes villes de l'espace d'étude, et pour un nombre réduit d'équipements (notamment les cafés-restaurants, banques, pharmacies). L'évaluation de la qualité d'OSM doit aussi tenir compte du caractère évolutif des bases contributives et du fort processus de rectification qui accompagne la production des données. On peut ainsi se demander dans quelle mesure l'évolution des contributions à OSM tendra à rééquilibrer ou à renforcer les contrastes observés à ces échelles infra et internationales. Il ne s'agit pas non plus d'attribuer à OSM un rôle (se substituer aux bases institutionnelles) que le projet ne prétend pas porter (Hayat, 2019, p.159), même si plusieurs travaux scientifiques ont montré que les informations contenues dans OSM atteignaient parfois des niveaux de qualité proches de ceux des bases conventionnelles.

Dans l'immédiat, un prolongement possible consisterait à reproduire cette méthode aux équipements de la BPE pour d'autres régions françaises, afin de mieux évaluer les facteurs de variation de ces résultats. Les premières pistes de comparaison internationale devraient par ailleurs être complétées et consolidées en élargissant la sélection à d'autres zones d'enquêtes allemandes et luxembourgeoises, caractéristiques d'autres contextes spatiaux (notamment les centres-villes ou zones commerciales de plus grandes agglomérations).

Bibliographie

Bakillah, M., Liang, S. H. L., Mobasheri, A., Jokar Arsanjani, J., Zipf, A.,(2014). Fine resolution population mapping using OpenStreetMap points-of-interest. International Journal of Geographical Information Science (IJGIS). Taylor & Francis. doi:10.1080/13658816.2014.909045

Barczak, A., Hilal M. 2016. “L ’Accès Aux Commerces et Services Dans Les Territoires de Vie Du Quotidien.” In *Campagnes Contemporaines. Enjeux économiques et Sociaux Des Espaces Ruraux Français*, edited by Stéphane Blancard, 97–116.

Baudet-Michel S., Conti B., Commenges H., Delage M., Guérois M., Guiraud N., Le Neindre C., Madry P., Paulus F., Pavard A., Toutin G., Vallée J., 2019, *La rétraction des services et commerces dans les villes petites et moyennes : modalités et logiques à l’échelle interurbaine*, rapport auprès du CGET et de la Caisse des dépôts, 206 p.

Devillet G., Jaspard M., Vazquez Parras J., 2012, « L’offre transfrontalière en commerce de détail », *Atlas interactif de la "Grande Région SaarLorLux"*, https://gr-atlas.uni.lu/index.php/fr/articles/wi55/gr1195/texte-integral

Griffond-Boitier, A. 1995. “L ’Apparition Des équipements Sportifs Sur Le Territoire National : L’identification Des Facteurs Structurants et Leurs Effets.” In *Sport, Relations Sociales et Action Collective*, edited by Jean-Paul (dir.) Augustin Jean-Pierre; Callède, 465–76.

Guérois, M., Ysebaert R., Giraud T., Madelin M., Louvet R., Michel Y., Feredj A., Grasland C. 2014. “Etude de Faisabilité d’une Observation Pérenne Des Espaces Transfrontaliers.” CGET - UMS RIATE.

Guérois, M., Ysebaert R., Viry M., Giraud T., Lambert N., Feredj A., Hamez G., De Ruffray S.. 2016. “Typologie Socio-économique Des Régions Frontalières de L’Union Européenne (2000-2012).” CGET - UMS RIATE.

Haklay, M. How good is volunteered geographic information? A comparative study of OpenStreetMap and ordnance survey datasets. Environ. Plan. B **2010**, 37, 682–703.

Hamez, G., Sophie de Ruffray, Claude Grasland, Nicolas Lambert, Amandine Hamm, and Emilie Moron. 2011. “Enjeux Des Territoires Frontaliers à L’échelle Nationale : Vers La Détermination et La Délimitation de Pôles Transfrontaliers.” CGET - UMS RIATE. https://doi.org/10.13140/2.1.4160.0646.

Hamez, G., Sophie de Ruffray, Florence Smits, Danielle Meddahi, and Emilie Moron. 2008. “La Comparaison Des Territoires Transfrontaliers à L’échelle Nationale : Vers Une Typologie Fondée Sur Les Logiques d’intégration.” CGET - UMS RIATE. https://doi.org/10.13140/2.1.3611.2004.

Hamez G., de Ruffray S., 2009, « La dimension sociale de la cohésion territoriale. L’exemple de l’accessibilité aux maternités dans la Grande Région », *Espace Géographique*, 2009-4, pp.328-344.

Hayat F., 2019, *Production des biens communs numériques et usages cartographiques*, Université Paris Diderot, 370 p.

Hombiat A., Villanova-Oliver M., Gensel J., 2016, « Un métamodèle pour structurer la folksonomie d’OSM en une nouvelle ontologie », *Revue internationale de géomatique* – n° 4/2016, 425-444.

Goodchild, M.F. (2007). "Citizens as sensors: the world of volunteered geography". GeoJournal. **69** (4): 211–221.

INSEE, 2018, *Base Permanente des Equipements 2018*, 16 p

Mashadi, A.; Quattrone, G.; Capra, L. The impact of society on volunteered geographic information: The case of OpenStreetMap. In OpenStreetMap in GIScience; Springer International Publishing: Cham, Switzerland, 2015; pp. 125–141.

Openshaw, S.. 1981. “Le Problème de L’agrégation Spatiale En Géographie.” *L’espace Géographique* n° 10-1: 15–24.

Petibon, C.. 2000. “Du Café Au Cinéma, La Hiérarchie Des équipements Selon Leur Seuil d’apparition.” *INSEE Midi-Pyrénées, 6 Pages* n°42: 1–6.

Stewart, J. Q. 1942. “A Measure of the Influence of a Population at a Distance.” *Sociometry* n° 5: 63–71. https://doi.org/10.4000/books.msha.16187.

Smits, F. 2006. “Les Pratiques Commerciales Dans Une Région Frontalière : Les Exemples d’Arlon, de Luxembourg-Ville et de Metz.” *Belgeo* 1-2: 163–78. https://doi.org/https://doi.org/10.4000/belgeo.11100.

Touya G, Antoniou V, Olteanu-Raimond AM, Van Damme MD. Assessing crowdsourced POI quality: Combining methods based on reference data, history, and spatial relations. ISPRS International Journal of Geo-Information. 2017; 6(3):80. https://doi.org/10.3390/ijgi6030080]

Vienne, F., Douay, N., Le Goix, R. and Severo, M. (2017) Les territoires du réseau social facebook : le cas des pratiques de géoréférencements, Territoire en mouvement Revue de géographie et aménagement, 34

Viry, M., Timothée Giraud, Marianne Guérois, Timothée Giraud, Ysebaert Ronan, Nicolas Lambert, and Amel Feredj. 2016. “Généralités Liées à Openstreetmap et à La Complétude Des Données).” CGET - UMS RIATE.

Ysebaert, R., Matthieu Viry, Marianne Guérois, Timothée Giraud, Lambert Nicolas, Valérie Lapenne, Gilles Toutin, and Vincent Nicat. 2018. “Regioviz, Un Nouvel Outil de Visualisation et de Comparaison Régionale).” CGET - UMS RIATE.