



HAL
open science

Apports et limites d'OpenStreetMap pour l'analyse spatiale des équipements commerciaux en zone transfrontalière - Note de synthèse

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 d'OpenStreetMap pour l'analyse spatiale des équipements commerciaux en zone transfrontalière - Note de synthèse: une étude de cas dans la Grande région. [Rapport de recherche] RIATE. 2020, 5 p. hal-03587237

HAL Id: hal-03587237

<https://hal.science/hal-03587237>

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 et limites d'OpenStreetMap pour l'analyse spatiale des équipements commerciaux en zone transfrontalière : une étude de cas dans la Grande région.

Synthèse étude ANCT – UMS RIATE – Octobre 2020

Marianne Guérois, Ronan Ysebaert, Timothée Giraud, Béatrice Maranget (Université de Paris, CNRS, RIATE)
Grégory Hamez, Mathias Boquet, Nicolas Dorkel (Université de Lorraine, LOTERR)

Introduction

L'objectif de cette étude méthodologique est de documenter l'apport potentiel de la base de données contributive *OpenStreetMap* (OSM) pour analyser l'organisation spatiale des commerces dans les espaces transfrontaliers. Alors que l'observation des centralités commerciales à l'échelle transfrontalière recouvre d'importants enjeux liés à l'érosion des équipements des petites villes, aux inégalités d'accessibilité qui en découlent, ou encore aux situations de concurrence et de complémentarité accentuées par les nouvelles opportunités d'accès aux équipements des Etats voisins, l'analyse statistique et cartographique de ces phénomènes se heurte à l'obstacle majeur de la disponibilité de données comparables : il n'existe pas en effet d'équivalent à la Base Permanente des Equipements (BPE) de l'INSEE au-delà des frontières, ces lacunes étant partiellement comblées par la réalisation d'enquêtes sur de petites portions de territoires ou par des études portant sur un type d'équipement au prix d'un long travail d'harmonisation des bases. **Dans ce contexte, la base OSM pourrait constituer une ressource précieuse du fait des informations collectées sur les « points d'intérêt » (POI, pour *Points Of Interest*)**. Dans la littérature, l'intérêt de recourir à OSM dans une perspective d'analyse territoriale a surtout été discuté à propos des réseaux d'infrastructure routière. De nombreux travaux se sont interrogés sur la qualité de données produites selon un mode décentralisé et non supervisé. Certains auteurs ont ainsi montré qu'au-delà d'un certain seuil de contributions, les jeux de données issus d'OSM sont comparables aux données traditionnelles maintenues par les agences nationales et que leur degré de complétude augmente avec le temps, ces constats étant relativisés par l'hétérogénéité spatiale de la complétude, plus élevée dans les agglomérations et dans les Etats où les communautés OSM sont les plus actives.

Qu'en est-il des données sur les équipements commerciaux ?

Dans quelle mesure le recours aux POI d'OSM permet-il de compenser l'absence d'une base de référence équivalente à la BPE de l'INSEE dans les Etats voisins ? Quelle est l'ampleur des lacunes de cette base contributive et dans quelle mesure sa complétude varie-t-elle selon les types d'équipements et le contexte de peuplement ? Peut-on identifier des pôles de concentration d'équipements sur la base d'OSM, pour toute la zone

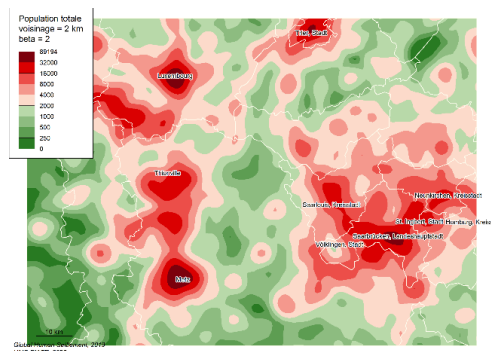


Figure 2 – Un espace d'étude transfrontalier, entre France, Allemagne et Luxembourg

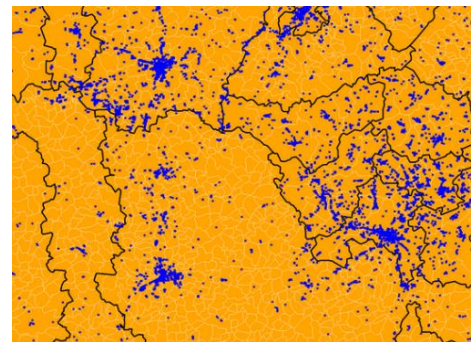


Figure 1 - Les Points d'Intérêt d'OSM pour analyser les équipements commerciaux à l'échelle transfrontalière ? (sélection de 15 équipements)

transfrontalière, en partant des équipements les mieux décrits ?

L'étude propose un éclairage sur ces questions en partant du cas d'un espace transfrontalier compris entre les villes de Luxembourg au Nord-Ouest, Metz au Sud, Sarrebrück à l'Est, Trèves au Nord-est (Figure 2), qui offre une grande variété de situations d'encadrement territorial par les commerces. L'ensemble de la démarche (analyses et représentations graphiques) est présenté dans un site web qui reprend les programmes construits avec le logiciel R (<https://frama.link/centralite>), de manière à exposer précisément la méthodologie suivie, et

éventuellement pouvoir transposer les différentes étapes d'extraction, de traitement et de représentation des données à d'autres espaces d'étude.

Evaluer la complétude des POI d'OSM à l'aide de bases de données d'équipements et de population

La comparaison avec différentes sources de données a permis d'évaluer la complétude des données OSM, de manière plus ou moins directe. Du côté français, l'existence de la Base Permanente des Equipements (BPE, INSEE) offre une opportunité unique en Europe pour mener finement cette comparaison à une échelle régionale. A l'échelle transfrontalière, la démarche est plus complexe compte tenu de l'absence de bases institutionnelles équivalentes dans les Etats voisins. Deux sources complémentaires ont par conséquent été mobilisées : d'une part, des relevés de terrain des équipements dans plusieurs rues commerçantes de 6 localités (quatre en France, deux en Allemagne), d'autre part des données de population issues de la grille kilométrique distribuée par la Commission Européenne (*Global Human Settlement*), qui tiennent lieu de variable de contrôle en supposant que la densité d'équipements (de proximité notamment) est proportionnelle à la densité de la population résidente.

L'étude s'appuie sur une sélection de 15 équipements fréquemment mobilisés dans la littérature sur les centralités commerciales, depuis les commerces de proximité jusqu'à quelques équipements plus rares correspondant à la gamme supérieure de la BPE (Figure 3). Une des étapes importantes consiste à faire le lien entre des catégories d'équipements décrites selon des logiques nettement différentes dans la nomenclature de la BPE et dans OSM. Dans OSM, en effet, il n'existe pas de nomenclature à proprement parler, mais des objets décrits à l'aide de tags (« étiquettes ») composés de couples de « clé-valeur » (par exemple « amenity » = « bank ») dont la liste, documentée dans un Wiki, est ouverte et non contraignante. Une table de correspondance OSM/BPE a ainsi été créée pour reconstituer une cohérence sémantique entre les deux bases (Figure 4). Pour certains équipements, la correspondance est directe et simple (ex :

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

Figure 3 - Les 15 équipements retenus pour l'étude (nomenclature BPE)

boulangeries). Pour d'autres, elle est plus ambiguë et a nécessité un travail de reconstitution plus important à partir du Wiki d'OSM (ex : alimentation). La majorité des catégories d'intérêt de la BPE ont ainsi pu être associées à des clés/valeurs d'OSM, à l'exception de certaines d'entre elles (les grandes surfaces alimentaires par exemple).

Équipement	Gamme	BPE (2018)	Tag OSM (clé/valeur)
Alimentation	Proximité	B201/ B202	shop=convenience /shop=greengrocer /shop=alcohol / shop=farm /shop=dairy / shop=cheese /shop=frozen_food
Ameublement	Intermédiaire	B306	shop=furniture /shop=antiques / shop=bathroom_furnishing /shop= lamps /shop=kitchen /shop=bed
Banque	Intermédiaire	A203	amenity=bank

Figure 4 - Extrait de la table des correspondances entre les catégories BPE et les tags OSM

Dans les communes françaises, des équipements OSM trois fois moins nombreux que dans la BPE, une complétude plutôt élevée dans les communes de centres-villes

Le premier volet de l'analyse de complétude porte sur les communes françaises et sur la comparaison entre OSM et la BPE, pour les 15 équipements sélectionnés : dans l'ensemble, les équipements OSM sont environ 3 fois moins nombreux que dans la BPE et il y a 2 fois plus de communes non équipées

dans OSM. Néanmoins, ce degré de complétude varie sensiblement selon les contextes spatiaux (Figure 5) : plus élevé 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, suivant une logique spatiale plus complexe qu'un simple gradient centre-périphérie. C'est ce que soulignent aussi les résultats sur la présence/absence des équipements, certaines zones urbaines semblant largement sous-contribuées, malgré leur forte densité de population et d'équipement (c'est le cas par exemple des communes-centres de Hayange et Sarreguemines). Par ailleurs, la situation varie fortement d'un équipement à l'autre (Figure 6). 8 équipements sont mieux contribués que la moyenne : 2 d'entre eux, théâtres et stations-services, sont plus nombreux dans OSM que dans la BPE du fait d'une définition nettement moins restrictive. D'autres présentent une bonne complétude relative (entre 40% et 90%), dont on peut supposer qu'elle est liée aux usages dominants d'OSM : cafés-restaurants, magasins d'alimentation, banques, pharmacies, magasins d'électronique et cinémas.

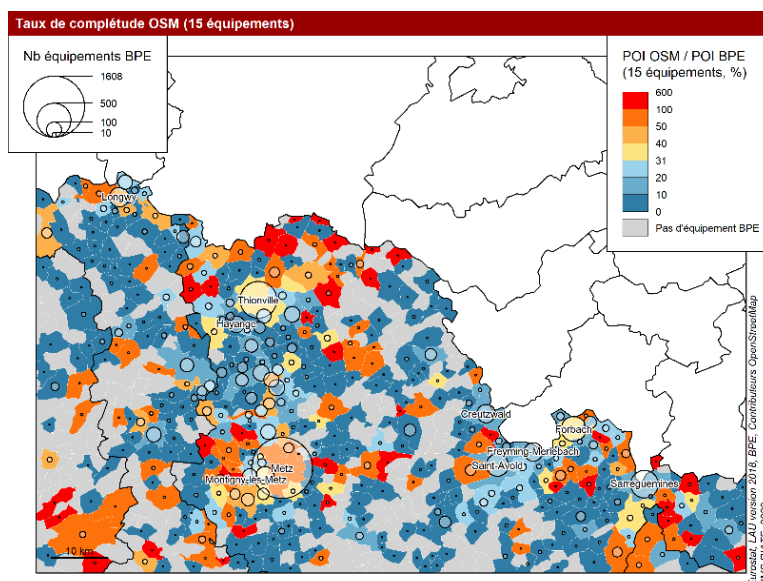


Figure 5 - Taux de complétude OSM par rapport à la BPE (15 équipements)

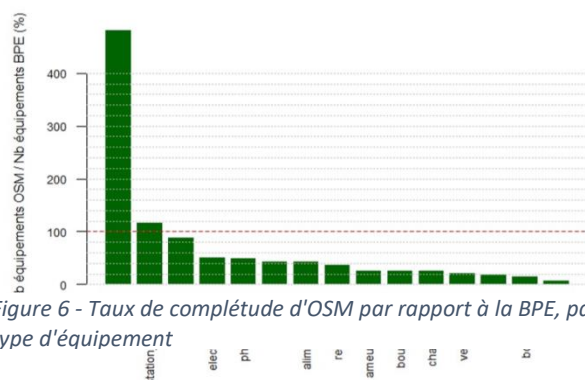


Figure 6 - Taux de complétude d'OSM par rapport à la BPE, par type d'équipement

communes françaises. Ils reflètent différents niveaux hiérarchiques d'équipement et présentent de nombreuses similarités selon les deux sources, 70% des communes étant associées au même profil d'équipement dans OSM et la BPE. Les données OSM conduisent toutefois à surestimer l'absence d'équipements, dans des communes qui sont la plupart du temps de petites communes rurales hors aires urbaines, mais qui sont pour une quinzaine d'entre elles des communes de banlieue (par exemple à Metz, Longwy, Sarreguemines) associées dans la BPE à un niveau intermédiaire d'équipements (présence d'au moins une banque et une pharmacie...).

En partant d'une information sur la présence/absence des 6 équipements les mieux contribués dans OSM (Figure 7), 5 profils communaux ont pu être identifiés pour les 703 communes françaises.

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)
4	Cinéma, parfois magasin d'électronique (15 communes) Ex : Longwy, Amnéville, St-Avold, Woippy (tout sauf un théâtre)... mais aussi qqs communes de 2000 habitants (Conflans en Jarnisy, banlieue de Jarny, possède tous équipements -y compris cinéma- sauf banque et théâtre. 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'agglomération de Jœuf, 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 7 - Niveaux d'équipements présents selon la BPE et OSM (résultats d'une Analyse en Composantes Multiples suivie d'une Classification Ascendante Hiérarchique)

A l'échelle transfrontalière : une méthode pour identifier les centralités d'équipements à partir d'OSM, une interprétation limitée par l'hétérogénéité internationale et locale des contributions

En l'absence de bases équivalentes à la BPE en Allemagne et au Luxembourg, l'analyse de complétude ne peut pas être étendue telle quelle à tout l'espace d'étude. Toutefois, plusieurs éléments de comparaison entre OSM et d'autres sources de données apportent un éclairage utile à la question de la validité de ces constats au-delà des frontières. L'enquête de terrain menée dans deux petites villes allemandes frontalières montre que les données OSM y sont nettement mieux renseignées que dans les villes françaises voisines, en accord avec la littérature soulignant l'existence d'une communauté OSM plus ancienne et plus active en Allemagne. De manière plus générale, les bonnes contributions relatives en Allemagne et au Luxembourg ressortent très nettement lorsqu'on évalue l'intensité de la relation entre potentiels locaux de population et d'équipements (ici les cafés-restaurants, banques, pharmacies et magasins d'électronique) et que l'on représente les écarts à cette relation (Figures 8 et 9) : ces écarts peuvent être interprétés en termes de sur-contribution relative à OSM lorsque le nombre

d'équipements est plus important qu'attendu compte tenu du nombre d'habitants (en bleu), et de sous-contribution dans le cas inverse (en rouge). Mise à part la zone environnant Metz et l'extrême sud-ouest de l'espace d'étude, ces cartes font ressortir l'existence d'une très nette

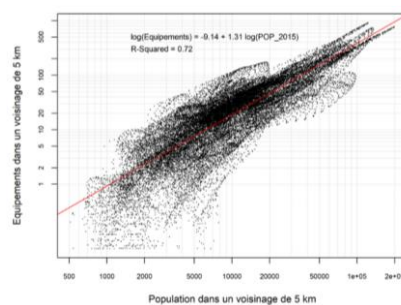


Figure 8 - Une relation étroite entre potentiel de population et potentiel d'équipements OSM

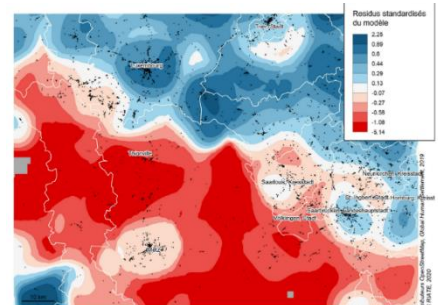


Figure 9 - Ecart à la relation entre potentiel de population et potentiel d'équipements OSM dans un voisinage de 5 km

discontinuité le long de la frontière, avec moins d'équipements OSM que ne le laisserait supposer la population côté français.

Tout en interprétant les résultats avec prudence du fait de l'hétérogénéité internationale et locale des contributions, nous proposons une représentation des centralités d'équipements qui s'appuie sur l'identification des zones de plus forte concentration de commerces par rapport à leur environnement proche (Figure 10). Ces « pics » relatifs correspondent aux plus forts différentiels entre un potentiel

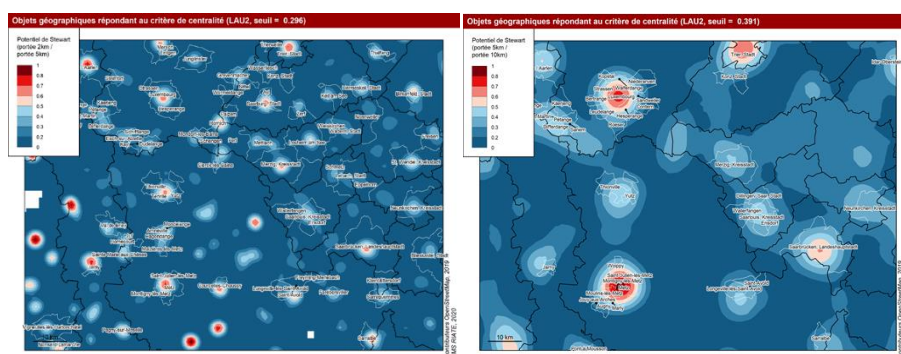


Figure 10 - Pôles locaux d'équipements : rapports de potentiels entre 2 et 5 km (à gauche) et 5 et 10 km de voisinage (à droite)

observé à proximité immédiate d'un lieu (par exemple à 2 km) et le même potentiel dans un voisinage élargi (5 km). Les noms de municipalités associées à ces « pics » sont issus du croisement avec les unités administratives sous-jacentes. Le choix de portées de voisinages plus ou

moins étendues permet d'ajuster la représentation des concentrations d'équipements à différents échelons d'observation : la carte de droite, par exemple, fait bien 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 d'équipements plutôt concentrées en France et dispersées en Allemagne.

Conclusion

Cette étude montre l'exploitation prudente qui peut être faite d'OSM pour travailler sur les équipements commerciaux en zone transfrontalière. **L'information 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 en France par rapport aux pays voisins. **En l'état actuel, la base OSM ne peut donc compenser l'absence de base nationale de référence sur les équipements en Allemagne et au Luxembourg, sauf pour les zones centrales des plus grandes villes de l'espace d'étude, et pour un nombre réduit d'équipements.** Toutefois, 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. Sans pour autant attribuer à OSM un rôle de substitution aux bases institutionnelles que le projet ne prétend pas porter, plusieurs travaux scientifiques ont montré que les informations contenues dans OSM atteignaient parfois des niveaux de qualité proches de ceux des bases conventionnelles. **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.** Dans le court terme, il serait utile de consolider les premières pistes de comparaison internationale en élargissant la sélection des lieux d'enquête à d'autres zones en Allemagne et au Luxembourg, plus caractéristiques des centres-villes et des zones commerciales de grandes agglomérations. Par ailleurs, l'analyse de la complétude d'OSM par rapport à la BPE pourrait aisément être étendue à d'autres régions françaises, en repartant de la démarche documentée dans le site internet qui supporte la méthodologie déployée pour cette étude (Figure 11), ceci afin non seulement de mieux évaluer les facteurs de variation de la qualité des contributions dans OSM, mais aussi de préciser les éléments de complémentarité entre les deux bases.

```
library(sf)
bbox <- st_read("data/geom.gpkg", layer = "bbox", quiet = TRUE)
lau50 <- st_read("data/geom.gpkg", layer = "lau50", quiet = TRUE)
lau <- st_read("data/geom.gpkg", layer = "lau", quiet = TRUE)
nuts3 <- st_read("data/geom.gpkg", layer = "nuts3", quiet = TRUE)
grid <- st_read("data/geom.gpkg", layer = "grid", quiet = TRUE)
#correctv <- st_read("data/osm.osm", layer = "correctv", quiet = TRUE)
```

Figure 11 - Une démarche reproductible : extrait du site internet commentant les étapes de la chaîne de traitements dans R (<https://frama.link/centralite>)

Bibliographie sélective (sur la qualité d'OSM et la méthode de lissage par potentiels)

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

Hamez, G., de Ruffray S., Grasland C., Lambert N., Hamm A., Moron E., 2011, *Enjeux des territoires frontaliers à l'échelle nationale : vers la détermination et la délimitation de Pôles transfrontaliers*. Rapport CGET - UMS RIATE.

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

Stewart, J. Q., 1942, "A Measure of the Influence of a Population at a Distance." *Sociometry* n° 5: 63–71.

Touya G., Antoniou V., Olteanu-Raimond A.M., Van Damme M.D., 2017, "Assessing crowdsourced POI quality: Combining methods based on reference data, history, and spatial relations", *ISPRS International Journal of Geo-Information*; 6(3):80.

Viry, M., Giraud T., Guérois M., Ysebaert R., Lambert N., Feredj A., 2016. "Généralités Liées à Openstreetmap et à La Complétude Des Données" CGET - UMS RIATE.