

Caractérisation du tissu économique local

Rapprochement entre le répertoire SIRENE et les Fichiers Fonciers



Credit photo : @ Manuel Bouquet - Terra

Sommaire

Sommaire

Problématique et intérêt de l'étude.....	4
L'enrichissement de la base Sirene par les Fichiers fonciers.....	4
Antériorité.....	4
Déroulement de l'étude.....	4
Intérêt de l'étude.....	4
Contenu et limites des Fichiers fonciers et du répertoire Sirene.....	5
Les Fichiers fonciers.....	5
Le répertoire Sirene.....	5
Points d'attention relatif au répertoire siren.....	6
Le rapprochement entre Sirene et les Fichiers fonciers.....	8
Problématique générale.....	8
Rapprochement au local ou à la parcelle ?.....	9
Résultats.....	10
Le rendu est réalisé à la TUP.....	10
Résultats globaux.....	10
Résultats sur la commune de Rambouillet.....	12
Procédure de test.....	15
Conclusion.....	17
Le contenu de l'étude.....	17
Limites.....	17
Intérêt et apports de l'étude.....	18
Annexe 1 : Méthode de rapprochement entre le répertoire Sirene et les parcelles des Fichiers fonciers.....	19
Méthode générale.....	19
Étapes d'affectation.....	19
Possibilité d'évolution de la méthodologie.....	20
Temps de traitement.....	20
Annexe 2 : Forme du rendu.....	21
Type de rendu.....	21
Annexe 3 : retour du service DADDT de la DRIEA.....	22
Annexe 4 : Présence d'établissements et de locaux d'activités sur le même emplacement.....	28
Annexe 5 : Lien entre les codes NAF dans les Fichiers fonciers et la base Sirene.....	30

Bordereau Documentaire

Informations du document

Titre : **Caractérisation du tissu économique local – Rapprochement entre le répertoire SIRENE et les Fichiers Fonciers**

Date du document : 24 octobre 2018

Diffusion Confidentiel (diffusion réservée au Cerema)
 Diffusion restreinte
 Diffusion libre

Auteur(s)

Martin Bocquet

Chargé d'études aménagement urbanisme foncier – rédacteur

Relecture Qualité

Frédéric Lasseron – Responsable du Groupe

Organisme(s) Auteur(s)

Nom de l'organisme : Cerema Nord-Picardie

Adresse 44 Ter, rue Jean BART CS 20275 59019 LILLE

Numéro de téléphone : 03 20 49 62 71

Adresse mail : martin.bocquet@cerema.fr

Organisme commanditaire

Nom de l'organisme Direction Régionale et Interdépartementale de l'Équipement et de l'Aménagement d'Île-de-France (DRIEA)



Adresse : 21-23 Rue Miollis, 75015 Paris

Informations contractuelles

Nature du rapport Intermédiaire
 Définitif

Numéro d'affaire (SIGMA) : C17NR0244

Visas techniques

Le chargé d'affaire : Bocquet Martin	Le responsable de groupe : Frédéric Lasseron
	

Historique des versions

Version	Date	Commentaire
V1	24 octobre 2018	Version 1 du rapport, accompagné des données
V2	29 novembre 2018	Version 2 du rapport et modification des données suites aux remarques de la DRIEA

Problématique et intérêt de l'étude

L'enrichissement de la base Sirene par les Fichiers fonciers

La DRIEA réalise actuellement des études sur l'implantation des activités (localisation des nouvelles activités, étude énergétique des bâtiments...). Dans ce cadre, elle souhaite disposer des Fichiers fonciers, enrichis par les données de la base Sirene.

Dans ce cas, la jointure se fera à l'adresse, afin de récupérer les informations sur l'occupant. En effet, le propriétaire est souvent une activité de holding, une banque ou un investisseur, dont l'identité ne permet pas de répondre aux besoins de la DRIEA.

Antériorité

Ce croisement a déjà été réalisé sur le département du Nord, avec des résultats satisfaisants pour les activités commerciales (environ 80 % de réussite sur la jointure). Cependant, cette méthode n'a pas encore été appliquée à un autre territoire.

L'enjeu est de savoir si cette méthode peut être exportée avec un taux de réussite suffisant, et si ce taux de jointure peut être amélioré.

Déroulement de l'étude

Cette étude s'est déroulée en plusieurs phases :

- constitution et échange sur la méthodologie,
 - test de la méthodologie sur un certain nombre de communes,
 - correction de la méthode,
 - production des données sur l'Île-de-France.
-

Intérêt de l'étude

Les Fichiers fonciers contiennent entre autres les éléments liés au bâti, notamment économique. De leur côté, la base de données Sirene contient les éléments liés au tissu économique et aux entreprises.

Ainsi, les Fichiers fonciers contiennent des informations relatives au bâti d'un commerce (étage, surface, implantation, nom du propriétaire...), et les fichiers Sirene les informations sur l'occupant (type d'entreprise, date de création, secteur d'activité...).

Il serait ainsi possible de rapprocher les données relatives au tissu économique de celles relatives au bâtiment, afin de comprendre les conditions d'occupation et les évolutions du type d'occupant sur un secteur particulier.

Contenu et limites des Fichiers fonciers et du répertoire Sirene

Les Fichiers fonciers	Les Fichiers fonciers sont une base de données utilisée pour le recouvrement de la taxe foncière. Ils contiennent donc toutes les informations sur le propriétaire du bien, ainsi que les caractéristiques des parcelles et des locaux ¹ .
Contenu des Fichiers fonciers	Deux tables seront utilisées dans le cadre de cette étude : <ul style="list-style-type: none">• la table des parcelles reprend toutes les informations liées aux parcelles,• la table des locaux reprend toutes les informations liées à un local, à usage d'activité ou d'habitation Par exemple, un bâtiment est présent sur cette parcelle. Un commerce est situé au rez-de-chaussée, et 4 logements au premier étage, pour un total de 5 locaux. Il est donc plus pertinent de récupérer les données au local. Cependant, le répertoire Sirene ne dispose que de l'adresse de la parcelle. L'affectation à un local particulier est donc complexe.
Millésime utilisé	On utilise le millésime 2017 des Fichiers fonciers, correspondant aux données au 1 ^{er} janvier 2017.
Limites dans le cadre de l'étude	Dans le cadre de l'étude, on peut dénombrer plusieurs limites ou points d'attention : <ul style="list-style-type: none">• les Fichiers fonciers contiennent peu d'informations liées aux propriétaires publics. Dans ce cadre, les établissements présents dans un local appartenant à un acteur public peuvent ne pas être repérés. Il s'agit d'un point d'attention important.• Les Fichiers fonciers ne contiennent pas de données liées au non-cadastré. Dans ce contexte, certaines échoppes sur le domaine public (ex : bords de fleuve...) peuvent ne pas être présentes. Ce point d'attention est cependant peu limitant.
Le répertoire Sirene	Le répertoire Sirene correspond à la liste des établissements français. En déclarant la création d'une entreprise (c'est-à-dire d'un établissement au sens de Sirene), celui-ci est inscrit directement dans le répertoire Sirene.
Code siren ou code siret ?	Il s'agit ici de faire le lien entre le code siret (c'est-à-dire code siren + code établissement) et le numéro de parcelle. Dans le reste du rapport, on parlera par abus de langage du code siren pour parler du siret.
Contenu de la base Sirene	La base Sirene contient plusieurs types d'informations liées à l'établissement, et donc entre autres : <ul style="list-style-type: none">• le code siren et siret,• des variables d'adresse,• la raison sociale de l'entreprise,• sa nature juridique,• son activité (selon la nomenclature NAF),• son caractère productif ou non,• l'année d'introduction dans la base.

¹Le site <https://datafoncier.cerema.fr/> regroupe les informations liées aux Fichiers fonciers.

Points d'attention relatif au répertoire siren

Lors de la création d'une entreprise, il est nécessaire de la déclarer au registre du commerce et des sociétés, ce qui entraîne son apparition de l'établissement associé dans le répertoire Sirene. Les entreprises sont ainsi en théorie toutes présentes dans la base. Il faut toutefois noter qu'il existe une procédure individuelle permettant aux entrepreneurs individuels de ne plus figurer dans ce fichier.

On peut donc considérer que les informations de cette base sont quasi-exhaustives. À l'inverse, les conditions de radiation du répertoire Sirene sont mal connues, et on retrouve encore dans le répertoire certains établissements clairement fermés.

Attention : ces points d'attention sont présentés à titre d'illustration, et ne se veulent pas être une expertise complète sur les limites du répertoire Sirene.

Entreprises inactives

En dehors des établissements clairement fermés subsistant encore dans le fichier, on retrouve dans le répertoire de nombreux établissements peu ou pas actifs. Il s'agit généralement d'établissements unipersonnels ou d'associations. Par exemple :

- un artiste occasionnel a constitué un établissement pour vendre ponctuellement plusieurs œuvres. Cet établissement est un appoint à son activité principale. Il n'a ainsi plus vendu d'œuvres depuis 5 ans. Il garde cependant cet établissement « au cas où » une occasion se représenterait.
- une association a été constituée et a déclaré une partie d'activité lucrative. Elle n'est cependant plus active, sans être officiellement dissoute.
- Un caviste décide de lancer une activité de type « dégustation ». Ne sachant pas si cette activité pouvait ou non se développer, il a créé un deuxième établissement, située au même endroit, afin de séparer ses activités. Quelques mois après, l'activité est intégrée à son activité principale de vente. Cependant, son « établissement de test » reste dans le répertoire, faute d'une démarche active du caviste pour la fermer.

Ces établissements inactifs constituent un point d'attention très important. Dans notre étude, nous souhaitons ainsi caractériser les établissements actifs, constitutifs d'un tissu économique local. La présence de ces établissements inactifs constitue ainsi un frein important à notre traitement.

SCI et copropriétés

Le répertoire Sirene contient en outre des données liées aux SCI et aux copropriétés. Ces établissements peuvent être nombreuses. Cependant, elles ne participent pas forcément à la constitution d'un tissu économique local.

Tri sur les activités

Pour notre étude, nous tentons d'étudier le tissu économique local. Dans ce cadre, nous excluons toutes les établissements qui n'ont pas de salariés². Ce tri permet ainsi d'exclure les SCI / copropriétés, ainsi qu'une partie des établissements inactifs.

De même, nous excluons les établissements dont la nature juridique pourrait correspondre à une SCI ou une copropriété³.

Mise à jour des données

Les données Sirene peuvent comporter des problèmes de mise à jour. En l'occurrence, les données liées au nombre de salariés ont l'air d'être peu mises à jour. Une expertise complémentaire serait ainsi nécessaire pour évaluer la qualité et les conditions de mise à jour de la donnée Sirene.

En particulier, les conditions de mise à jour de l'adresse ne sont aujourd'hui pas connues. Les tests réalisés semblent néanmoins conclure à la fiabilité de la donnée.

²C'est-à-dire les établissements dont l'effectif est différent des modalités « 00 » et « NN » .

³C'est-à-dire différent de '6540' (Société civile immobilière), '6599' (Autre société civile) ou '9110' (Syndicat de copropriété).

Données utilisées Le répertoire Sirene est en open data depuis le 1^{er} janvier 2017. Les données sont actualisées en continu. Cela signifie notamment que les données anciennes sont écrasées au fur et à mesure.

Le répertoire existe sous deux formes :

- le répertoire littéral, fourni en open data. **Il est à noter que deux formes de fichiers coexistent actuellement jusqu'à janvier 2019 : l'étude est ainsi basée sur l'ancienne structure de données.**
- l'association OpenStreetMap géolocalise les adresses de ce répertoire en continu.

Dans les deux cas, il est nécessaire d'indiquer la date de validité des données.

Pour cette étude nous utilisons les données littérales au 1^{er} janvier 2017. Cependant, pour améliorer le rapprochement, nous utilisons en outre les données géolocalisées. Faute d'un téléchargement récent, les données les plus anciennes disponibles sont celles au 1^{er} mai 2018. On reprend ainsi les données de géolocalisation que pour les établissements existants à la fois au 1^{er} janvier 2017 et au 1^{er} janvier 2018. L'objectif est donc de rapprocher les données littérales Sirene au 1^{er} janvier 2017 avec les Fichiers fonciers au 1^{er} janvier 2017.

Le rapprochement entre Sirene et les Fichiers fonciers

Problématique générale

Il s'agit ici de faire le lien entre :

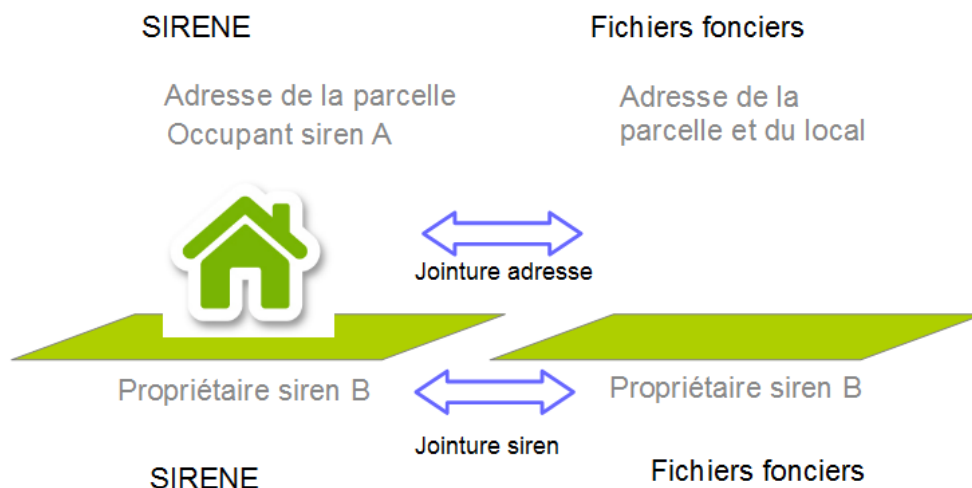
- les données relatives aux locaux tels qu'inscrits dans les Fichiers fonciers,
- les données relatives à l'occupation économique tels qu'inscrites dans le répertoire Sirene.

La dissociation propriétaire / locataire pour les locaux économiques

Les Fichiers fonciers contiennent des données relatives au propriétaire et le fichier Sirene correspond aux données de l'occupant. Par exemple, le commerce « PUDDING ET NAPPERON » est locataire de sa boutique appartenant à « MISS MARPLE SA ». Les Fichiers fonciers contiendront les informations liées au local (surface, nombre de dépendances, localisation de la parcelle, etc.) et les informations liées au propriétaire « MARGOULIN SA ». Les fichiers Sirene contiendront les informations liées à l'entreprise « PUDDING ET NAPPERON » (lieu du siège social, nombre d'employés, type d'activité, etc.).

Il n'y a donc aucun lien simple entre ces données. Il est donc nécessaire de passer par une jointure entre l'adresse du bien (adresse de la parcelle ou du local dans les Fichiers fonciers) et l'adresse de l'entreprise (adresse de l'entreprise dans Sirene).

Variables communes



Le rapprochement entre les Fichiers fonciers et le répertoire Sirene se fait sur le numéro Sirene si l'on souhaite faire un rapprochement du propriétaire, et à l'adresse si l'on souhaite faire un rapprochement de l'occupant.

Source : Cerema Nord-Picardie - Magali Journet

Problème de normalisation des adresses

Cette jointure à l'adresse est possible, mais pose un certain nombre de problèmes.

En premier lieu, les adresses sont remplies, dans les deux bases, par une déclaration des occupants, dont la forme n'est pas normalisée. En d'autres termes, les adresses suivantes portent sur le même bien :

- 1, rue du général de Gaulle,
- 1 r du general de Gaulle,
- 1 rue grl de Gaulle.

Le premier exercice est donc d'arriver à normaliser les adresses.

Ce problème est notamment très important dans les zones d'activités ou les zones commerciales.

Rapprochement au local ou à la parcelle ?

L'étude vise à produire deux rapprochements différents :

- un rapprochement entre une parcelle et un code siren. On sait ainsi que le commerce « ALLO TUYAU » est situé sur la parcelle 001.
- un rapprochement entre un local et un code siren. Une fois que l'on connaît la parcelle sur laquelle le code siren est situé, on tente d'assigner le code siren à un local en particulier. Dans le cas précédent, on saura ainsi que le commerce est situé au rez-de-chaussée, et a une surface déclarée de 70 m².

Questions liées à la jointure

La jointure entre l'adresse du répertoire Sirene et l'adresse dans les Fichiers fonciers est complexe. En effet :

- à une parcelle peut correspondre plusieurs codes siren : il s'agit par exemple de deux bureaux dans un même immeuble.
- À un local peut correspondre plusieurs codes siren. Un bureau est par exemple loué à deux avocats qui exercent chacun leur activité : on aura ainsi deux codes siren pour un local.
- À l'inverse, un code siren peut correspondre à plusieurs parcelles, si celles-ci ont la même adresse.

De manière générale, on peut ainsi se trouver dans une relation « N-N », c'est-à-dire qu'à plusieurs codes siren correspondent plusieurs locaux, voire plusieurs parcelles.

Abandon de la jointure au local

Face à une jointure au local peu satisfaisante, seule la jointure à la parcelle a été réalisée. Les éléments qui ont guidé ce choix sont explicités dans la partie « résultats ». Cependant, quelques pistes de méthodes possibles, qui n'ont pas été mises en œuvre dans cette étude, sont présentées.

Résultats

Le rendu est réalisé à la TUP

Les Fichiers fonciers ont certains biais, notamment dus aux problématiques d'unités foncière et de copropriétés multiparcellaire : si deux parcelles adjacentes ont le même propriétaire, les données liées au bâti sont souvent concentrées sur une seule parcelle⁴. Dans ce cadre, le Cerema reconstitue des entités fictives, les « TUP », qui constituent le regroupement des parcelles adjacentes appartenant au même propriétaire. Il s'agit en d'autres termes de la plus petite unité sur laquelle les informations sont fiables.⁵

Forme du rendu

Le rendu et les calculs sont réalisés à la TUP. Dans ce cadre, on traite bien du rapprochement des TUP avec chaque entreprise. **Dans la suite du rapport, par abus de langage, on utilisera parfois le terme de « parcelle » pour parler des TUP.**

Résultats globaux

Au niveau de la région Île-de-France, le taux de jointure global est de 91,1 %. Cela signifie que 91,1 % des entreprises listées dans la base Sirene trouvent une affectation sur une ou plusieurs TUP.

La méthode est présentée en annexe.

Différence selon les types d'espaces

De manière générale, le rapprochement est plutôt bon dans les zones les plus urbaines, et un peu moins bon dans les zones urbaines et périurbaines. Cependant, c'est au niveau des entreprises situées dans les zones d'activités et les centres commerciaux que se situent la plupart des problèmes. Dans de telles zones, le taux de géolocalisation chute de manière importante.

Différence selon les départements

Tous les départements n'ont pas le même taux de géolocalisation. En pratique, la géolocalisation est bien meilleure dans les départements plus urbains et possédant peu de complexes d'activité (zones d'activités, zones commerciales...), que dans les départements de banlieue.

Département	Taux de jointure
75	97,3 %
77	77,8 %
78	84,0 %
91	83,2 %
92	95,0 %
93	91,3 %
94	91,0 %
95	84,1 %

Taux d'entreprises dans la base Sirene pour lesquelles on trouve une correspondance dans les Fichiers fonciers

⁴Par exemple, les parcelles A et B appartiennent au même propriétaire. Une maison est construite sur chaque parcelle. Dans ce cas, les données peuvent recenser 2 maisons sur la parcelle A et 0 sur la parcelle B.

⁵Plus d'informations sont disponibles dans le guide de prise en main des Fichiers fonciers, disponible à cette adresse : https://datafoncier.cerema.fr/ressources-ff/prise-en-main#bien_demarrer

Affectation selon l'étape Au niveau régional, la plupart des parcelles sont affectées lors des premières étapes (égalité parfaite des adresses après retraitement). L'utilisation de Sirene géolocalisée par OpenStreetMap permet de réaliser la jointure pour 45 000 lignes supplémentaires, faisant passer le taux de jointure de 80 à 90 % environ.

Etape d'affectation	Nombre de locaux affectés	% de locaux affectés lors de l'étape	Pourcentage cumulé
Etape 2 : liaison entre l'adresse et la table des parcelles	283 096	58,8%	58,8%
Etape 3 : liaison entre l'adresse et les locaux	107 675	22,4%	81,2%
Etape 4 : liaison sur un autre champ adresse sirene et la table des parcelles	2 489	0,5%	81,7%
Etape 5 : liaison sur un autre champ adresse sirene et la table des locaux	1 719	0,4%	82,1%
Etape 6 : liaison avec Sirene géolocalisée	42 392	8,8%	90,9%
Etape 7 : Jointure particulière pour les centres commerciaux	1 044	0,2%	91,1%
Non affecté	42 974	8,9%	100,0%
Total Résultat	481 389	100,00%	

Affectation selon l'étape (méthodologie et étapes présentées en annexe)

Typologie des locaux non affectés A la fin du traitement, différents établissements restent non affectés. Ils ont généralement une des caractéristiques suivantes :

- une adresse mal remplie dans la base Sirene (absence de numéro, rue mal orthographiée...),
 - une adresse située sur le domaine public, ou lié à un local public (donc pas forcément renseigné dans les Fichiers fonciers),
 - leur présence dans des zones d'activités, des lieux-dits, des centres commerciaux.
- D'après les tests de terrain réalisés, le premier point est majoritaire.

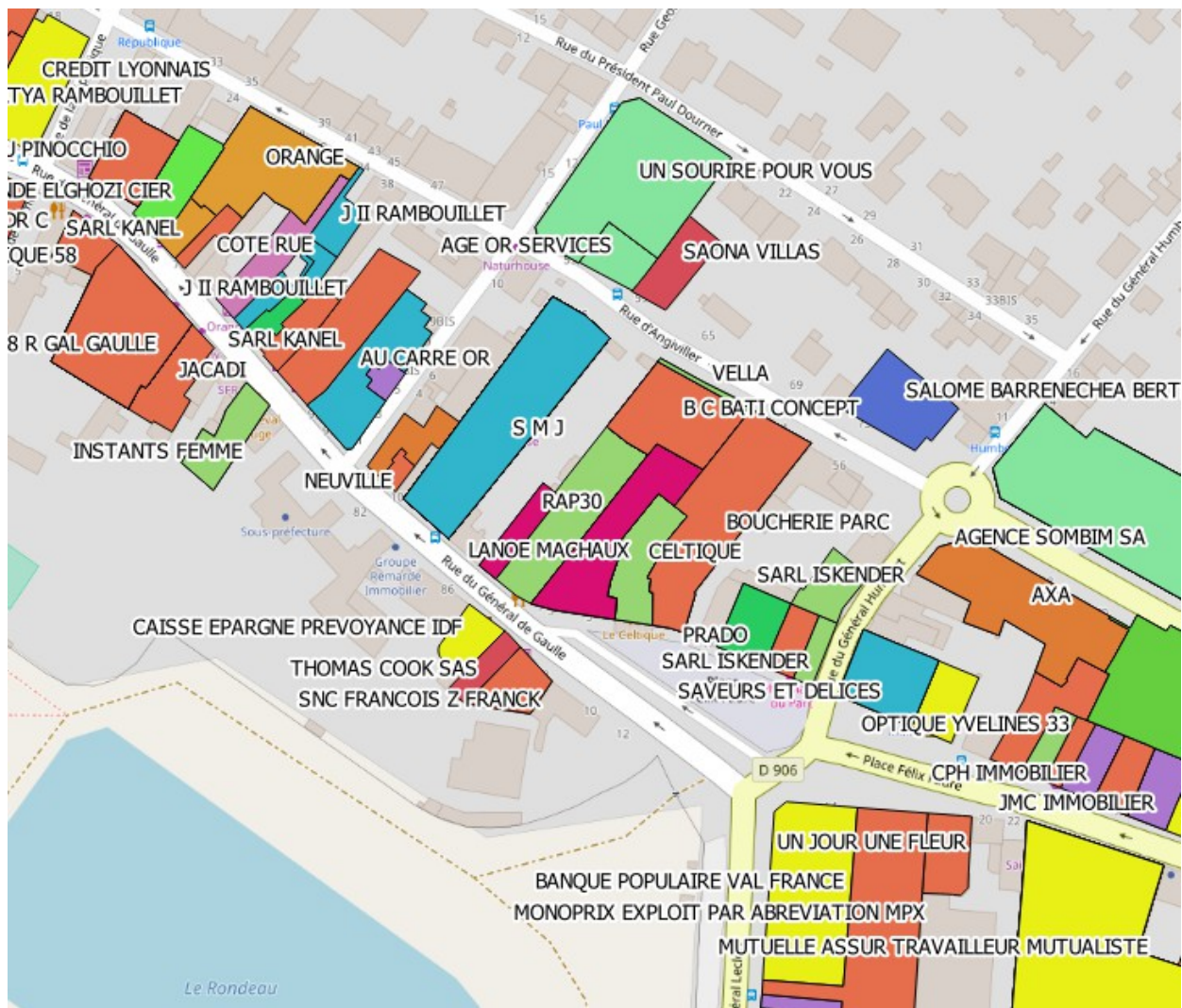
Résultats sur la commune de Rambouillet

Test sur le centre-ville

Cette étude permet d'observer le tissu économique à une maille très fine (rue, centre-ville...). En particulier, la variable intéressante est ici le code NAF de l'établissement, ainsi que sa raison sociale.

Si l'on réalise un test sur le centre-ville, on obtient la carte suivante. Il faut cependant noter les précautions suivantes prises dans la présentation de la carte :

- plusieurs établissements sont présentes sur la même parcelle. Dans ce cas, seule l'une d'entre elles a été présentée. Par exemple, la parcelle dont le nom « SNJ » apparaît ci-dessous abrite en réalité 13 établissements dont les noms ne sont pas affichés sur la carte.
- Un tri a été réalisé pour ne pas afficher les établissements avec un nom de particulier (« Monsieur Jean Dupond »).
- Le nom de l'établissement peut être significativement différent de l'enseigne. Par exemple, la boutique avec la devanture « Coiffure C'Hair Emma », pourra en réalité être un établissement inscrit au nom de « Emma Coiffure ». Le nom «Emma Coiffure » sera donc affiché sur la carte.



Centre-ville de Rambouillet, avec les noms des différents établissements. Source : Sirene / Fichiers fonciers 2017. Les entités en couleurs sont les TUP sur lesquelles on trouve un établissement avec au moins un salarié.

La carte ci-dessous indique en couleur les catégories NAF correspondantes⁶. Par exemple, les parcelles en orange affichent les parcelles sur lesquels il y a un local commercial⁷.



Centre-ville de Rambouillet. Les TUP en orange ont au moins un commerce embauchant au moins un salarié.

Test sur le centre commercial

La zone commerciale de Rambouillet est composée de plusieurs parcelles. En particulier, la grande parcelle contient le centre commercial « Belair ». On retrouve sur cette emprise (grande parcelle) 38 établissements répertoriés dans Sirene.

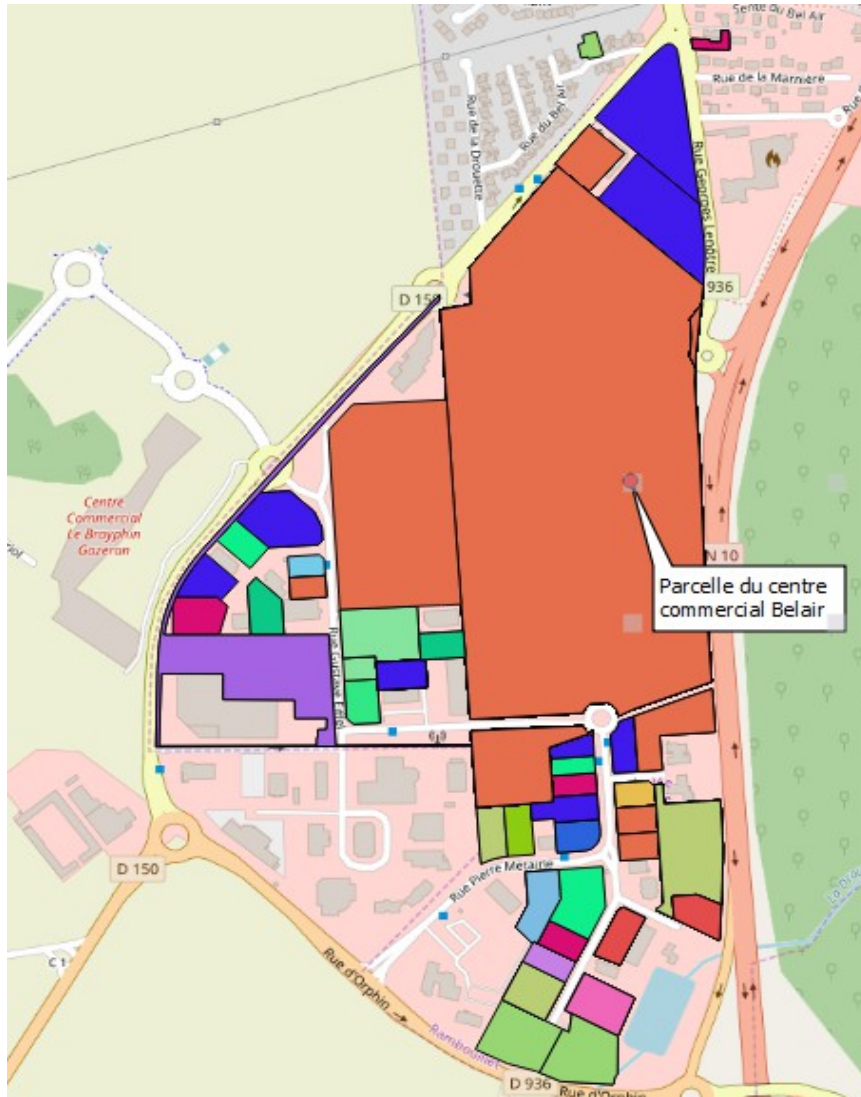
Les établissements géolocalisés l'ont été avec succès : il n'y a donc pas de faux positif (établissement mal localisée à cet endroit). Il faut cependant noter que seules 38 établissements sont présentes. A l'inverse, le centre commercial est censé héberger 50 établissements différentes⁸.

Cet exemple est assez emblématique du traitement des zones commerciales : les boutiques situées à cet endroit sont crédibles, mais le recensement n'est pas exhaustif. De même, sur certains autres centres, il peut subsister des problèmes de géolocalisation.

⁶ Il s'agit ici des « divisions » au sens de la NAF : <https://www.insee.fr/fr/information/2120875>

⁷Avec les mêmes limites que précédemment : si plusieurs établissements sont présentes sur une parcelle, seule l'activité d'une d'entre elles sera affichée

⁸En faisant le décompte des enseignes présentes sur le site internet du centre (<http://belair-rambouillet.com/>). Il est à noter que ce site revendique pour sa part 60 boutiques.



Zone commerciale de Rambouillet

Procédure de test

Le test de la méthode s'est porté sur 9 communes :

Nom de la commune	Code INSEE	Taux de géolocalisation à la tup	Nombre de code siren totaux
Courbevoie	92026	95%	3903
Puteaux	92062	84%	3180
nanterre	92050	91%	3317
Gentilly	94037	92%	581
Rungis	94065	63%	1082
Paris 14	75014	98%	7985
Serris	77449	88%	687
Bray-sur-Seine	77051	71%	121
Montereau-Fault-Yonne	77305	87%	665

Ces envois ont fait l'objet de deux commentaires complets, transcrits en annexe.

Question de l'adresse dans les Fichiers fonciers

Les Fichiers fonciers contiennent des informations liées à l'adresse du bien (parcelle ou local). Les tests effectués montrent cependant des cas particuliers pour lesquels l'adresse telle qu'inscrite dans les Fichiers fonciers n'est pas bonne.

Il s'agit notamment de cas de parcelles présentes dans des ensembles (notamment La Défense).

Présence de deux établissements avec le même code Sirene sur une parcelle

Il peut y avoir certains cas pour lesquels deux établissements (siret différents), mais avec le même code Sirene (c'est-à-dire la même entreprise) cohabitent sur une même parcelle.

Après vérification, il s'agit d'un cas courant visant à séparer les activités d'une même entreprise, et donc à noter dans des comptes à part deux activités dans des établissements différents. Il peut s'agir par exemple d'un menuisier ayant une activité existante de création de charpentes. Afin de se diversifier, il fait l'acquisition de nouvelles machines destinées à réaliser des meubles. Pour des raisons fiscales, et éviter de faire peser les risques sur son premier établissement, il crée un nouvel établissement pour son activité de menuiserie.

Il ne s'agit généralement pas d'une erreur, mais bien du choix du chef d'entreprise de séparer ses activités.

Effectifs des établissements

Les premiers tests effectués ont permis de confirmer que les effectifs tels que renseignés dans la base Sirene étaient relativement peu mis à jour, et donc peu pertinents.

Les établissements sont-ils situés dans des locaux commerciaux ?

Une méthode de test a été réalisée est présentée dans l'annexe 4.

Ce premier exemple rappelle la différence fondamentale entre le local (dans le sens d'entité bâtie), telle que vue par les services fiscaux, et l'établissement. Cette différence est à l'origine d'un certain nombre de problèmes relatif au rapprochement entre les bases.

Les codes NAF Sirene et Fichiers fonciers correspondent-ils ?

Une méthode de test a été réalisée et est présentée dans l'annexe 5.

En conclusion, nous avons une convergence des codes NAF présents dans les données Sirene et les Fichiers fonciers. Cependant, à un niveau plus global, ces données sont issues de traitements et de logiques différentes, qui aboutissent in fine à une divergence de chiffres. Sur les établissements plus classiques (un établissement dans un local), on peut considérer que le code NAF correspond. Sur les établissements moins classiques (plusieurs établissements dans un local, établissements qui ne sont pas situés dans un local commercial...), il n'y a pas convergence.

***Géolocalisation des
TUP et rapprochement***

Il est rappelé que certaines parcelles des Fichiers fonciers ne sont pas géolocalisées. Ce cas est rare, mais peut ainsi concerner certaines parcelles des zones commerciales, comme c'est le cas dans l'annexe 3b.

***Géolocalisation de la
base Sirene par
OpenStreetMap***

De même, dans la base Sirene géolocalisée, OpenStreetMap a fait le choix de géolocaliser la totalité des établissements. A ce titre, si l'adresse est insuffisamment renseignée, l'établissement a été situé au milieu de la rue (si seul le numéro est manquant) ou au centre de la commune (si l'adresse n'a pas été reconnue).

Il a été fait le choix de ne pas tenir compte de ces géolocalisations approximatives dans le traitement.

Conclusion

Le contenu de l'étude

L'étude a donc permis de rapprocher les Fichiers fonciers, disposant de données sur le bâti à usage d'activité et le répertoire Sirene, qui reprend la liste et les caractéristiques de chaque établissement français.

Il s'agit de faire le lien entre la forme du bâti (ce local est un commerce) et les établissements (ces établissements sont présentes sur cette parcelle). C'est donc un rapprochement entre la parcelle et l'occupant du local (et non son propriétaire, dont les informations sont déjà présentes dans les Fichiers fonciers). Le lien entre les deux bases est ainsi fait à partir des adresses inscrites dans celles-ci.

Limites

Les différentes techniques de rapprochement ont permis d'atteindre un score de 91,1 % sur l'Île-de-France. Cela signifie que l'on peut faire le lien entre ces établissements et une ou plusieurs parcelles.

Ce rapprochement est de bonne facture en général. Les taux sont cependant plus bas dans les zones commerciales et les zones d'activités, pour lesquelles les adresses sont plus approximatives. Cependant, le problème principal tient au contenu des deux bases. Comme explicité précédemment, les Fichiers fonciers listent les locaux, et le répertoire Sirene les établissements. En pratique :

- certains locaux abritent plusieurs établissements. Il peut s'agir de la création d'une nouvelle activité en sus de celle déjà existante. Dans ce cadre, le gérant peut souhaiter, pour des raisons stratégiques et/ou fiscales, créer une autre établissements. Il peut en outre s'agir d'un local abritant plusieurs établissements indépendantes (par exemple plusieurs médecins, constituant autant d'établissements différentes, dans un même local).
- A l'inverse, un établissement peut posséder plusieurs locaux. Il peut ainsi s'agir de différents cas : des garages appartenant à l'établissement peuvent constituer des locaux différents, l'établissement possède plusieurs sites qui constituent autant de locaux...

On peut donc dire que les Fichiers fonciers traitent du tissu bâti sur lequel est hébergé une activité économique, et que le répertoire Sirene traite plutôt du tissu économique local, sans lien avec l'urbanisme.

Dans ce cadre, nous avons donc des logiques très différentes entre les bases, qui ne permettent pas de faire un rapprochement simple de type « 1 local =1 établissement».

D'autres bases de données pourraient permettre un meilleur rapprochement. En particulier, il peut être intéressant d'étudier le rapprochement de la base de données « CLAP » de l'INSEE⁹ avec les Fichiers fonciers.

⁹<https://www.insee.fr/fr/metadonnees/source/serie/s1162>

Intérêt et apports de l'étude

Le rapprochement entre les Fichiers fonciers, disposant de données sur le bâti à usage d'activité et le répertoire Sirene, qui reprend la liste et les caractéristiques de chaque établissement français est néanmoins intéressant sur plusieurs points :

- Il permet d'observer la totalité du tissu économique, en le localisant à l'échelle parcellaire.
- Il permet, à travers l'utilisation de plusieurs millésimes Sirene, de faire l'étude de l'apparition et de l'évolution des établissements, dans un périmètre particulier.
- Les logiques différentes des deux bases peuvent être combinées pour observer l'adéquation et l'évolution du potentiel bâti, et le mettre en relation avec la dynamique de création des établissements.
- Il permet de bénéficier rapidement d'une liste des établissements, pour réaliser des portraits économiques de territoires ou de quartiers.
- La typologie des activités, mesurée à travers les codes NAF des occupants, peut être mieux caractérisée à une échelle locale.
- Les Fichiers fonciers 2017 disposent de données plus fiables sur les locaux à usage d'activité. Dans ce cadre, la base Sirene permet de mieux connaître le tissu économique, tandis que les Fichiers fonciers se concentrent plutôt sur le bâti.

Annexe 1 : Méthode de rapprochement entre le répertoire Sirene et les parcelles des Fichiers fonciers

Méthode générale

Le rapprochement se fait en plusieurs étapes. Dans un premier temps, on normalise les adresses des deux bases, pour favoriser leur jointure. On essaye ensuite d'affecter les codes siren, avec des critères de moins en moins contraignants.

Il s'agit ainsi d'une méthode en tamis : un code Sirene déjà affecté ne sera plus modifié dans les étapes suivantes.

Étapes d'affectation

Étape 1 – nettoyage des adresses

On transforme les adresses issues des bases Sirene et des Fichiers fonciers afin de maximiser le taux de jointure. Dans ce cadre, les caractères spéciaux sont retirés des adresses. On retire des mots clés qui pouvaient s'avérer problématiques (LA, DE, etc.). certains mots-clés sont transformés (GNRL → GENERAL, PLACE → PL, etc.)

Étape 2 : liaison entre l'adresse et la table des parcelles

Dans un premier temps, on fait la jointure entre l'adresse Sirene et l'adresse de la parcelle dans les Fichiers fonciers. Seules les adresses correspondant parfaitement sont prises en compte.

Étape 3 : liaison entre l'adresse et les locaux

Dans les Fichiers fonciers, certains locaux ont une adresse différente de la parcelle qui les porte. On réalise ainsi la jointure entre l'adresse Sirene et les locaux. Cette liaison permet de récupérer le nom de la parcelle sur laquelle est situé le local.

Étapes 4 et 5 : liaison sur un autre champ adresse

Le répertoire Sirene contient plusieurs champs d'adresse. Pour maximiser le rapprochement, la jointure est ensuite effectuée sur un autre champ d'adresse (l4_normalisee).

L'étape 4 correspond à un rapprochement avec la table parcelle des Fichiers fonciers, l'étape 5 à un rapprochement avec la table des locaux.

Étape 6 : utilisation d'une base géolocalisée

La base Sirene a été géolocalisée par Openstreetmap. On utilise ainsi cette géolocalisation pour améliorer l'appariement de la base. Dans ce cas, on assigne à un code siren la parcelle la plus proche.

Sur cette étape, seules les adresses parfaitement géolocalisées sont prises en compte : on exclut ainsi les adresses ne possédant pas de numéro de rue.

Étape 7 : jointure approchée pour les centres commerciaux

Les zones d'activités et centres commerciaux ne disposent que rarement d'une adresse fiable. Dans ce contexte, pour les centres commerciaux, le rapprochement est fait sur l'ensemble du centre commercial, sans prendre en compte le numéro de rue.

Pour les adresses correspondant le mot clé « centre commercial », les étapes 2 à 5 sont ainsi répétées, sans tenir compte du numéro de rue.

Possibilité d'évolution de la méthodologie

Il existe plusieurs possibilités d'évolution.

On peut faire une jointure sur les noms des propriétaires, pour repérer les propriétaires occupants. Cependant, certains propriétaires possèdent plusieurs parcelles dans la commune. Dans ce cas, comment faire la différence entre chacun de ses biens ? Cette évolution n'a pas été mise en place aujourd'hui, notamment au regard du faible nombre d'affectation que cela occasionne, et du nombre important de mauvaises affectations.

Le redressement de l'étape 7 n'est aujourd'hui réalisée que pour les centres commerciaux. Dans ce cadre, il pourrait être intéressant de réaliser un redressement similaire pour les zones d'activités. Pour cela, une liste des noms de zones d'activités serait nécessaire.

Temps de traitement

Le temps de constitution de la table principale est assez faible. Suite à des évolutions de code, une commune est traitée en quelques secondes (de 2 secondes pour une petite commune à une dizaine de seconde pour une commune plus importante). Le traitement de la table principale (environ 500 000 établissements) à l'échelle de la région Île-de-France est ainsi de l'ordre de quelques heures.

La constitution des tables de rendu (création de la table finale à partir de la table de passage) est de l'ordre d'une heure environ.

Annexe 2 : Forme du rendu

Type de rendu

Le rendu est constitué de deux tables :

- une table au format .csv reprenant l'intégralité des données siren
- une table géolocalisée au format .sql.

Contenu des tables .csv

Les tables csv reprennent les données de la table Sirene, dont la description est disponible sur internet¹⁰. La structure de la base est identique (une ligne pour un établissement).

Deux champs ont été ajoutés :

- le champ idpar est de type tableau, et reprend toutes les parcelles pour lesquelles on a pu assigner une adresse. En d'autres termes, si l'adresse de l'établissement A correspond à la parcelle01 et la parcelle01, la colonne idpar sera égale à [Parcelle01, Parcelle02].
- Le champ « etape_affectation » trace l'étape (selon la typologie ci-dessus) à laquelle on a pu déterminer lier le code siren à la parcelle.

Les établissements dans lesquels la variable « idpar » est vide correspond aux établissements que l'on a pas pu affecter.

Contenu des tables .sql

Ces tables sont directement dérivées des tables précédentes. Elles réalisent la jonction entre les codes siren et les TUP.

Cette table est géolocalisée. En sus des champs précédents, elle contient la totalité des champs de la table des « TUP », dont la description est disponible sur le site internet des Fichiers fonciers¹¹.

La structure de la table est actuellement la suivante :

lorsque le code siren correspond à une seule TUP, le champ l_idpar correspond à la liste des parcelles présentes sur cette TUP

lorsque le code siren correspond à plusieurs TUP (ce qui est le cas lorsque une adresse correspond à plusieurs TUP), il y aura plusieurs lignes, correspondant à toutes les combinaisons possibles TUP – siren.

À titre d'exemple, si les établissements 1 et 2 correspondent aux TUP A et B, nous aurons 4 lignes dans la base :

- TUP A – établissement 1
- TUP B – établissement 1
- TUP A – établissement 2
- TUP B – établissement 2

¹⁰<https://www.sirene.fr/sirene/public/static/liste-variables>

¹¹<http://doc-datafoncier.cerema.fr/ff/docff/>

Annexe 3 : retour du service DADDT (C. Salagnac) de la DRIEA

Analyse du croisement des fichiers fonciers (MAJIC) et SIRENE (Insee) (sur la base des travaux du CEREMA)

1 / Champ

- Secteur d'activité : **commerce de détail**
- Secteur géographique : 2 communes « test » :
 - **Serris (77)** : implantation du centre commercial régional « Val d'Europe » et de « La Vallée Village » ;
 - **Bray-sur-Seine (77)** : connaissance fine du terrain (proche habitation).
- Millésimes des fichiers SIRENE OpenStreetMap (Insee) et Fichiers Fonciers (Dgfi) : 2016

2 / Comparaison de chiffres clés – Données de cadrage : Serris (77)

Croisement fichiers fonciers (surfaces en m²) / SIRENE (codes APET 700)

apet700	libapet	Nbre SIREN	Stotp (m ²)
4711B	Commerce d'alimentation générale	1	220
4711C	Supérettes	1	580
4711F	Hypermarchés	1	549
4719B	Autres commerces de détail en magasin non spécialisé	1	549
4724Z	Comm. détail pain pâtisserie & confiserie (magasin spécialisé)	6	15 385
4730Z	Commerce de détail de carburants en magasin spécialisé	1	549
4741Z	Comm. détail ordinateurs unités périph. & logiciels (magas. spéc)	1	549
4742Z	Comm. détail matériels télécommunication (magasin spécialisé)	4	557
4751Z	Commerce de détail de textiles en magasin spécialisé	5	33 588
4752A	Comm. détail de quincaillerie, peintures et verres (mag.< 400 m2)	3	1 841
4752B	Comm. détail de quincaillerie, peintures et verres (mag.> 400 m2)	1	549
4754Z	Commerce de détail appareils électroménagers (magasin spécialisé)	1	549
4759A	Commerce de détail de meubles	8	1 585
4759B	Commerce de détail d'autres équipements du foyer	8	41 864
4762Z	Commerce de détail de journaux et papeterie en magasin spécialisé	2	1 993
4764Z	Commerce de détail d'articles de sport en magasin spécialisé	4	1 367
4765Z	Commerce de détail de jeux et jouets en magasin spécialisé	3	346
4771Z	Commerce de détail d'habillement en magasin spécialisé	166	31 568
4772A	Commerce de détail de la chaussure	19	22 265
4772B	Commerce de détail de maroquinerie et d'articles de voyage	15	33 942
4773Z	Commerce de détail produits pharmaceutiques (magasin spécialisé)	2	565
4775Z	Comm. détail de parfumerie & produits de beauté en magasin spéc.	7	12 353
4776Z	Comm. dét. fleurs, plantes, etc, animaux de cie et leurs aliments	1	549
4777Z	Comm. détail d'articles horlogerie & bijouterie (magas. spéc.)	14	18 255
4778A	Commerces de détail d'optique	6	2 009
4778C	Autres commerces de détail spécialisés divers	7	24 079
4779Z	Commerce de détail de biens d'occasion en magasin	3	27 970
4791A	Vente à distance sur catalogue général	1	7 746
4791B	Vente à distance sur catalogue spécialisé	1	0
4799A	Vente à domicile	2	0
4799B	Vente par automate, aut. com. dét. hors mag., éventaire ou marché	1	3 821
TOTAL Commerce de détail > >		296	287 739
Codes APET autre que commerce de détail		446	-
Nombre total établissements SIRENE		742	-

stotp : Surface totale des parties professionnelles sur la parcelle (en m²)

Source : fichiers fonciers (DGFIP), Sirene (Insee) – Traitements : CEREMA – 2016

Commerce de plus de 300m²

typologie	nature	NB - siret
Culture sports loisirs	Culture/Loisirs	2
	Jouet	2
	Sport	1
	Téléphonie	1
Cycles auto	Centre auto	1
Équipement maison	Bricolage	1
	Cuisine/Salle de bain	1
	Electrodomestique	1
	Meuble	3
Équipement personne	Bijouterie	1
	Mode	21
Hygiène santé beauté	Beauté/Santé	2
Hypermarchés	Hypermarché	1
Restauration	Restauration	4
Supérettes	Supérette	1
Total >>		43

Source : Données issues du Visiaurif – IAU IdF – LSA 2014

Note méthodologique pour le calcul de stotp :
Dans le fichier initial (MAJIC/SIRENE), le tableur contient autant d'idtup que d'établissements sirene qui ont pu être joints via l'adresse. L'opération effectuée est donc la suivante : (somme stotp (GroupBy idtup and Apet700)) / nombre d'idtup correspondant.

ATTENTION : la surface « stotp » d'une parcelle renseigne sur la somme des surfaces de l'ensemble des locaux présents sur cette parcelle (quelle que soit leur nature). **L'indicateur surface renseigné ici n'est donc que très peu fiable**, car il suffit qu'un seul établissement SIRENE classé en commerce de détail soit présent sur une parcelle MAJIC pour que la surface totale de celle-ci soit renseignée. Mais cette parcelle contient, dans la plupart des cas, d'autres locaux. Il aurait fallu faire le lien avec la table des « locaux » dans MAJIC pour faire un rapprochement plus fin des établissements SIRENE et des locaux dans MAJIC afin d'apprécier leurs surfaces.

Commentaires :

En 2016, le croisement MAJIC / SIRENE a permis de comptabiliser **296 établissements du commerce de**

détail (au sens de la NAF Insee). **La source de données LSA comptabilise, quant à elle, 43 points de vente (commerces \geq 300 m²)** en 2014. LSA a catégorisé les points de vente selon une typologie spécifique qui ne permet pas de reprendre directement les codes NAF de l'INSEE. Néanmoins, certains rapprochements peuvent être effectués pour les appellations / terminologies identiques employées par les deux sources.

Par exemple :

- **Les supérettes / hypermarchés** : les deux sources ne recensent qu'une supérette et qu'un hypermarché ;
- **Jeux et jouets** : 3 établissements dans les fichiers SIRENE/MAJIC contre 2 dans la base LSA ;
- **Bijouteries** : 14 établissements dans SIRENE/MAJIC (code NAF4777Z) contre 1 dans la base LSA ;
- **Meubles** : 8 établissements dans SIRENE/MAJIC (code 4759A) et 3 dans la base LSA.

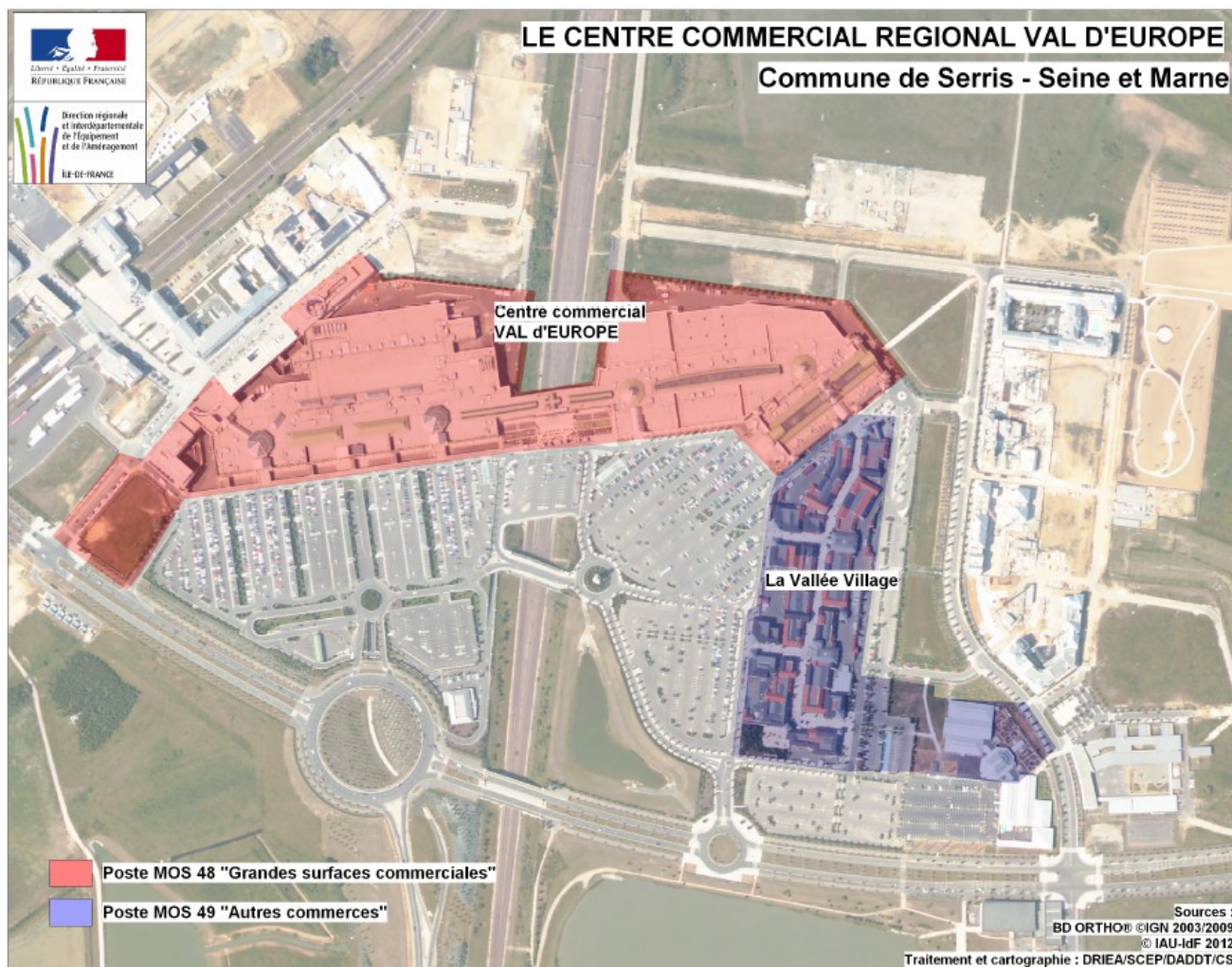
De gros écart sont à noter. Ces derniers peuvent notamment s'expliquer par :

- **La classification des commerces** (codes Insee pour les fichiers MAJIC / SIRENE) et la typologie propre à LSA (15 catégories) ;
- **Les millésimes** entre les deux sources sont différents : 2016 (MAJIC/SIRENE) et 2014 pour LSA ;
- **Le champ couvert** : tous les commerces dans MAJIC/SIRENE (sans distinction de taille) alors que dans la base LSA, seuls les commerces de plus de 300 m² sont comptabilisés ;
- Avant même d'effectuer le croisement entre les fichiers SIRENE/MAJIC, les données entre les deux sources peuvent varier sur les établissements du commerce de détail (exhaustivité des données, champ couvert, etc.).

3 / Le centre commercial régional « Val d'Europe » et la « Vallée Village »

La commune de Serris (77) est caractérisée par deux types de centres commerciaux :

- le centre commercial régional « Val d'Europe » (98 400 m²);
- le centre commercial spécialisé « La Vallée Village ».

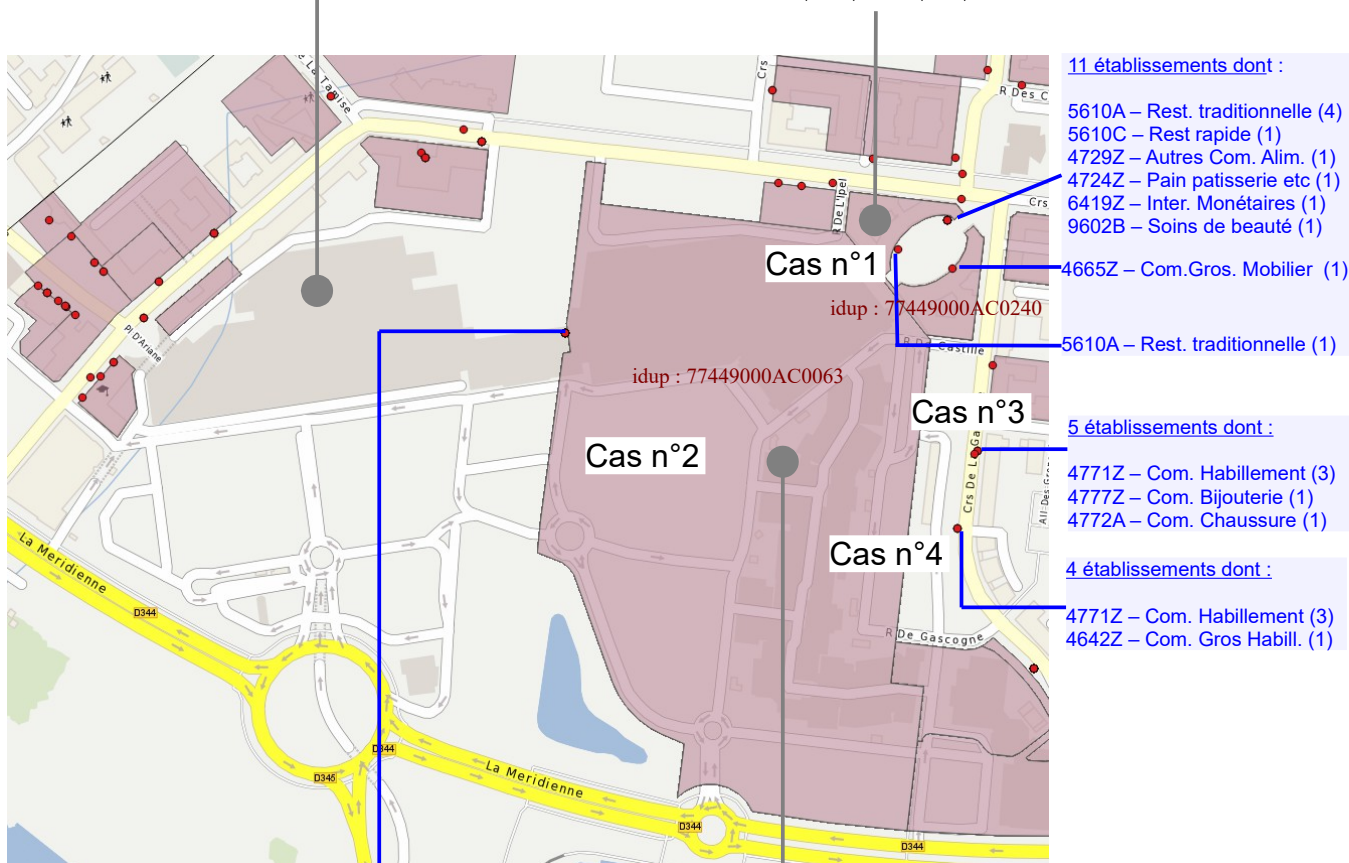


Remarque : le bâtiment (gris) figurant en fond de plan correspond au centre commercial « Val d'Europe » (voir carte page précédente). La partie gauche du centre commercial n'est donc pas couvert graphiquement par une parcelle TUP (Majic)

Informations contenues dans la TUP (jointure SIRENE / MAJIC)

apet700	libapet	NB - siren
4665Z	Commerce de gros de mobilier de bureau	1
4724Z	Comm. détail pain pâtisserie & confiserie (magasin	1
5610A	Restauration traditionnelle	4
5610C	Restauration de type rapide	1
6820B	Location de terrains et d'autres biens immobiliers	1
8130Z	Services d'aménagement paysager	1
Total Résultat >>		9

Source : fichiers fonciers (DGFIP), Sirene (Insee) – Traitements : CEREM



Informations contenues dans la TUP (jointure SIRENE / MAJIC)

apet700	libapet	NB - siren
4616Z	Interm. comm. textiles, habillt, fourrures, chaussures & art cui	1
4642Z	Commerce de gros d'habillement et de chaussures	3
4649Z	Commerce de gros d'autres biens domestiques	1
4724Z	Comm. détail pain pâtisserie & confiserie (magasin spécialisé)	1
4751Z	Commerce de détail de textiles en magasin spécialisé	2
4759B	Commerce de détail d'autres équipements du foyer	4
4771Z	Commerce de détail d'habillement en magasin spécialisé	62
4772A	Commerce de détail de la chaussure	5
4772B	Commerce de détail de maroquinerie et d'articles de voyage	6
4775Z	Comm. détail de parfumerie & produits de beauté en magasin	1
4777Z	Comm. détail d'articles horlogerie & bijouterie (magas. spéc.)	3
4778C	Autres commerces de détail spécialisés divers	2
4779Z	Commerce de détail de biens d'occasion en magasin	1
5610C	Restauration de type rapide	2
6820B	Location de terrains et d'autres biens immobiliers	1
7022Z	Conseil pour les affaires et autres conseils de gestion	1
Total Résultat		96

Source : fichiers fonciers (DGFIP), Sirene (Insee) – Traitements : CEREM – 2016

Remarque : cette parcelle couvre à la fois une partie du centre « Val d'Europe » et la quasi totalité de la « Vallée -Village »

Légende :

- Données Sirene localisées par OpenStreetMap (OSM)
- Parcelle TUP (Majic – Dgfp) Données issues du croisement SIRENE/MAJIC
- En bleu : données SIRENE OpenStreetMap

Commentaires

Cas n°1 : La place de Toscane

- Contexte : cette place est réputée pour ses nombreux points de restauration ;
- Le croisement SIRENE/MAJIC fait ressortir 9 établissements alors que les données SIRENE fournies par OpenStreetMap (OSM) en comptabilisent 11 sur le même périmètre ;
- Les deux sources comptabilisent le même nombre de commerces de restauration rapide (1), commerce de gros de bureau (1), pâtisserie / confiserie (1). La base SIRENE/OSM compte 1 point de vente de restauration traditionnelle de plus (5) que la base SIRENE/MAJIC (4) ;
- Sur les 11 établissements SIRENE recensés par OSM, on a donc :
 - 7 correspondances exactes avec la base SIRENE/MAJIC (Apet 700) ;
 - 2 établissements n'ayant pas de correspondance APET entre les deux sources ;
 - 2 établissements « manquants » dans la base SIRENE/MAJIC.

Cas n°2 : Le centre commercial Val d'Europe et la Vallée Village

- Remarque : la parcelle TUP (77449000AC0063) ne reprend que la moitié du périmètre du centre « Val d'Europe » et recouvre complètement le centre « la Vallée Village ». Dans les adresses des établissements compris dans cette parcelle, il est fait référence aux deux centres commerciaux ;
- La partie de gauche du centre commercial Val d'Europe qui n'est pas visuellement couverte par une TUP contient de nombreux commerces (connaissance terrain : « Paul », « Orange », « Bouygues Telecom », « Kiabi », « Marionaud », Coiffeurs, Cordonnerie, Restaurations Rapides, etc.). Ces commerces ne sont pas listés dans la TUP couvrant la partie droite du centre ;
- La TUP comptabilise 96 établissements (SIRENE/MAJIC), alors que la base SIRENE /OSM n'en liste que 12. Sur ces 96 établissements :
 - il y a le même nombre de commerces de restauration rapide (2) ;
 - il y a 10 fois plus de commerce de l'habillement dans la base MAJIC/SIRENE (62) que dans celle de SIRENE/OSM (6)
 - 4 établissements ont des codes Apet700 différents entre les deux sources ;
 - 84 établissements « manquants » dans la base SIRENE/OSM.

Cas n°3 : établissements localisés en milieu de la voie mais « basculés » dans une TUP

- 5 établissements SIRENE/OSM sont localisés (graphiquement) au milieu de la rue « Cours de la Garonne » ;
- Après vérification manuelle, ces 5 établissements ont tous été « basculés » dans la TUP voisine (77449000AC0063).

Cas n°4 : établissements localisés en milieu de la voie et non répertoriés dans une TUP

- 4 établissements SIRENE/OSM ont été placés en milieu de voie et reste à priori introuvables dans les TUP à proximité.

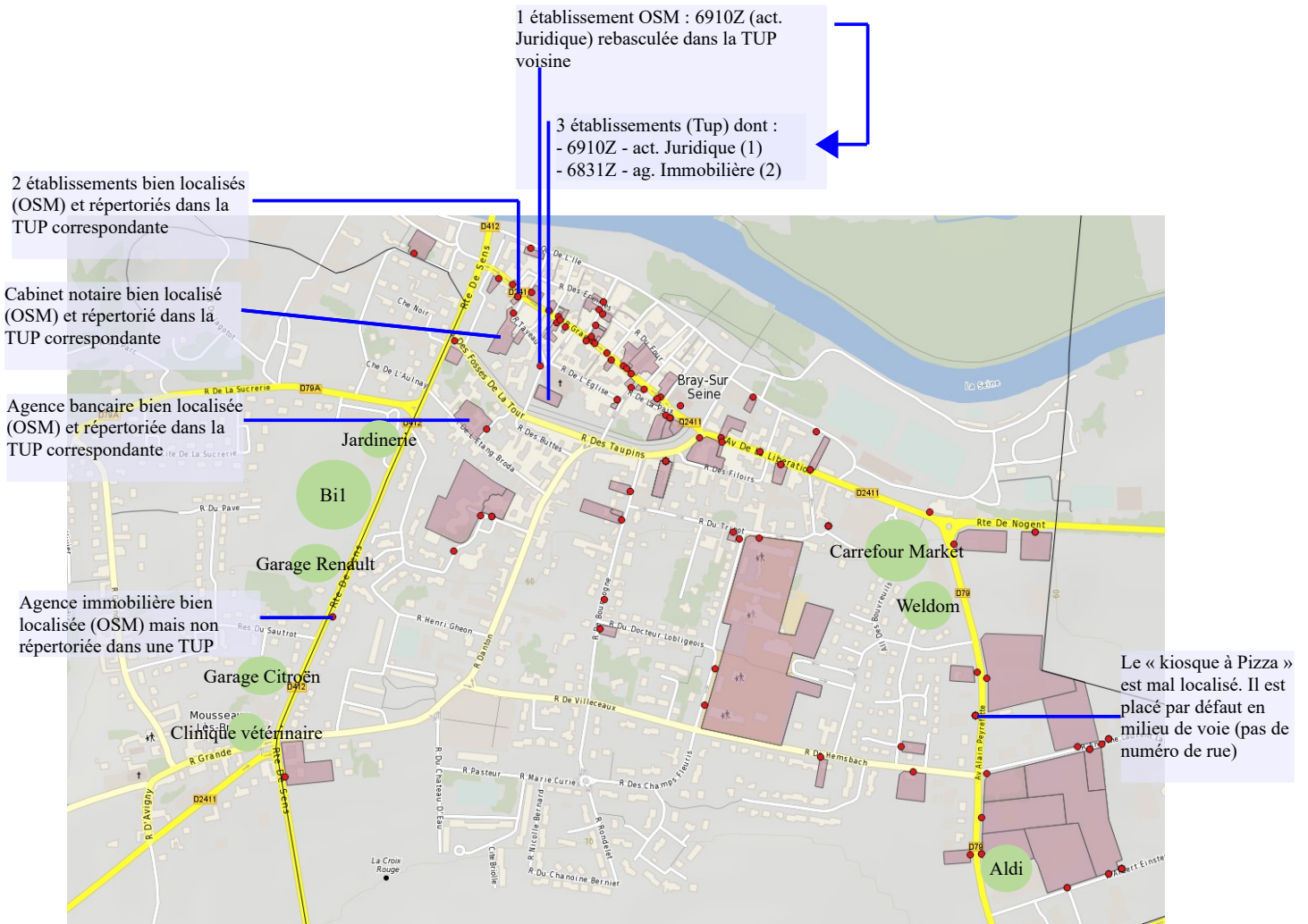
4 / Étude de cas : la commune de Bray sur Seine (77)

Contexte : connaissance du terrain, proche de mon lieu d'habitation.
Commune rurale de 2 200 habitants

Principaux constats : les principaux établissements du commerce (en surface) implantés dans la commune ne sont pas recensés ni dans la base SIRENE/OSM ni dans le rapprochement des fichiers SIRENE/MAJIC. **Plus globalement, il manque un certain nombre d'établissements commerciaux.** Quelques exemples

(recensement non exhaustif) figurent sur la carte ci-après. Leurs emplacements sont approximatifs mais illustrent les commerces manquants (d'après la connaissance terrain)

Cartographie – Bray sur Seine (77)

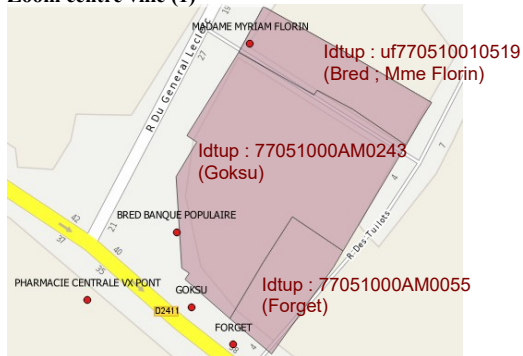


A titre d'illustration (surfaces de vente) :

- Carrefour Market : 1 900m²
- Aldi : 774 m²

Source : données issues du VisaairifIAU – Données LSA 2014

Zoom centre ville (1)



- On a 5 établissements SIRENE/OSM :
- Mme Florin (9602A – Coiffure)
 - Bred Banque Populaire (6419Z – Intermédiation monétaire)
 - Goksu (5610C – Rest. Rapide)
 - Forget (5630Z – Boissons)
 - Pharmacie centrale vieux pont (4773Z – Com Détail Pharmacies)

Cerema Nord-Picardie, DRIEA

→ On ne retrouve que 4 de ces établissements dans les Tup limitrophes (il manque la pharmacie)

Zoom centre ville (2)



Les établissements « Romairy » (5610A – Rest. Traditionnelle), SCBM (idem) et « Aubriot Rossignol Caumont » (Cabinet médical) sont bien localisés et repertoriés dans les TUP correspondantes.

Il manque 1 coiffeur à l'emplacement indiqué (connaissance terrain)

Zoom centre ville (3)



Les établissements « Mr Bricolage » et « Gedimat » sont bien localisés (OSM) et répertoriés dans les TUP correspondantes.

Il manque les établissements « Weldom » (magasin de bricolage) et « Carrefour Market » (supermarché)

Octobre 2018

Annexe 4 : Présence d'établissements et de locaux d'activités sur le même emplacement

Un établissement n'est pas forcément dans un local d'activité

On peut aussi comparer la présence de locaux d'activités sur une parcelle avec la présence d'un établissement. En d'autres termes, lorsque l'on dispose d'un local d'activité, avons nous aussi un seul code Sirene ? A l'inverse, lorsque l'on observe un code Sirene, trouvons nous un local d'activité ?

Les tests ont été réalisés sur le département 95. Les conclusions restent toutefois valides à l'échelle de l'Île-de-France.

Méthode Pour chaque TUP, on compte le nombre d'établissements issus de la base Sirene et le nombre de locaux d'activité issus des Fichiers fonciers. On compare ensuite, par catégorie, le nombre de locaux issus de chacune des deux bases.

Résultats On obtient ainsi les résultats suivants. Les tableaux notent en vertical le nombre de locaux dans recensés dans la base Sirene et en horizontal le nombre de locaux d'activités dans les Fichiers fonciers.

Par exemple, dans le département 95, il y a 1 743 TUP sur lesquelles on recense entre 2 et 5 établissements, mais un seul local d'activité. Cela correspond à 4 418 établissements au total (2^e tableau), et 1 743 locaux d'activités (3^e tableau).

Nombre de locaux sur la TUP selon les Fichiers fonciers

Nombre d'entreprises sur la TUP selon Sirene	Nombre de locaux sur la TUP selon les Fichiers fonciers					Total Résultat
	0	1	2-5	6-10	>10	
1	8 661	3 801	1 076	53	35	13 626
2-5	1 594	1 743	1 389	193	93	5 012
6-10	62	116	139	113	91	521
>10	44	64	68	38	138	352
Total Résultat	10 361	5 724	2 672	397	357	19 511

Nombre de locaux sur la TUP selon les Fichiers fonciers

Nombre d'entreprises sur la TUP selon Sirene	Nombre de locaux sur la TUP selon les Fichiers fonciers					Total Résultat
	0	1	2-5	6-10	>10	
1	8 661	3 801	1 076	53	35	13 626
2-5	3 918	4 418	3 842	649	338	13 165
6-10	458	855	1 032	852	724	3 921
>10	6 711	1 463	1 690	634	3 650	14 148
Total Résultat	19 748	10 537	7 640	2 188	4 747	44 860

Nombre de locaux sur la TUP selon les Fichiers fonciers

Nombre d'entreprises sur la TUP selon Sirene	Nombre de locaux sur la TUP selon les Fichiers fonciers					Total Résultat
	0	1	2-5	6-10	>10	
1	0	3 801	2 625	363	2 448	9 237
2-5	0	1 743	3 865	1 422	4 954	11 984
6-10	0	116	449	864	2 303	3 732
>10	0	64	213	295	6 921	7 493
Total Résultat	0	5 724	7 152	2 944	16 626	32 446

Comparaison du nombre de locaux selon les Fichiers fonciers et du nombre d'établissements recensés dans la base Sirene pour chaque parcelle. Les cases vertes sont celles pour lesquels les décomptes correspondent, et en rouge celles pour lesquelles les différences sont importantes

Analyse On constate de manière générale que les locaux ne correspondent pas aux établissements. En d'autres termes, la situation « classique », à savoir un commerce, correspondant à un local, n'est pas représentative.

En pratique :

- un local dans les Fichiers fonciers peut correspondre à plusieurs établissements. Il s'agit par exemple d'un commerce avec deux activités, que le propriétaire a décidé de séparer. Ou il peut s'agir par exemple de 3 avocats présents dans le même local, mais ayant chacun leur établissement.
- Une activité peut correspondre à plusieurs locaux dans les Fichiers fonciers. Il s'agit par exemple d'un commerce, pour lequel la surface de vente et le local de stockage est séparé. De même, un établissement peut posséder 1 local de bureaux et 19 locaux correspondant à 19 places de parking dans une copropriété. Nous aurons donc un établissement qui correspondra à 20 locaux d'activités.

Ce premier exemple rappelle la différence fondamentale entre le local (dans le sens d'entité bâtie), tel que défini par les services fiscaux et l'établissement. Cette différence est à l'origine d'un certain nombre de problèmes relatif au rapprochement entre les bases.

Annexe 5 : Lien entre les codes NAF dans les Fichiers fonciers et la base Sirene

Les codes NAF Sirene et Fichiers fonciers correspondent-ils ?

Les Fichiers fonciers contiennent, dans les données liées au local, le code NAF de l'occupant. Des tests ont ainsi été réalisés pour tenter de rapprocher le code NAF issu de Sirene avec le code NAF issu des Fichiers fonciers.

Limites

Ce rapprochement consiste en un décompte du nombre de codes NAF égaux sur chaque parcelle. Il faut cependant prendre en compte les limites suivantes :

- lorsqu'un établissement est géolocalisé sur plusieurs parcelles (ce qui est souvent le cas dans les TUP), le code Sirene est affecté à plusieurs parcelles. Par exemple, l'établissement A est située sur les parcelles A et B, et aura pour code NAF 4791B. Dans les Fichiers fonciers, le local est situé sur la parcelle A. EN termes de décompte, on aura donc 1 élément pour lequel il y aura correspondance, et une parcelle pour laquelle il n'y aura pas correspondance.
- Il y a une déconnexion importante entre les activités et les locaux occupés par ces dernières. En d'autres termes, un établissement peut occuper plusieurs locaux, et plusieurs établissements peuvent être présentes sur un local. Enfin, certains établissements ont l'adresse de leur siège différent de celle de leur activité réelle.
- Le Fichier Sirene contient toutes les établissements, y compris certains établissements peu actifs ou fermés. À l'inverse, les Fichiers fonciers ne recensent que une activité par local à vocation d'activité. Cela signifie que le nombre d'établissements classés dans Sirene est bien supérieur aux locaux d'activités présentes dans les Fichiers fonciers.

Divergence entre les bases

Une fois ces éléments écartés, on peut voir les divergences suivantes :

- de nombreux codes NAF sont présents dans Sirene, mais ne sont pas remplis dans les Fichiers fonciers. Cela est issu la plupart du temps du traitement qui a été réalisé, et qui localise le code Sirene à la TUP et non à la parcelle. Il faut donc relativiser le différentiel relativement important entre les deux chiffres.
- Il existe un certain nombre de codes NAF présents dans les Fichiers fonciers et pas dans Sirene. Il s'agit pour la plupart des données Sirene non géolocalisées, ou des établissements présents sur plusieurs locaux.

Convergences entre les bases

On peut cependant noter que lorsque des codes NAF existent des deux côtés, il y a souvent convergence.

En d'autres termes, on a donc convergence des codes NAF présents dans les données Sirene et les Fichiers fonciers. Cependant, à un niveau plus global, ces données sont issues de traitements et de logiques différentes, qui aboutissent in fine à une divergence de chiffres. Sur les établissements plus classiques (un établissement dans un local), on peut considérer que le code NAF correspond. Sur les établissements moins classiques (plusieurs établissements dans un local, établissements qui ne sont pas situés dans un local commercial...), il n'y a pas convergence.

Résultats

Dép	Nombre de codes NAF communs	Nombre de locaux possédant des codes sirene dans FF et dans Sirene	Nombre de codes NAF dans Sirene ne trouvant pas de correspondance dans FF	Nombre de codes NAF dans FF ne trouvant pas de correspondance dans Sirene
75	60 114	99 414	125 634	19 388
77	10 482	15 464	29 332	6 963
78	11 655	17 748	34 864	7 111
91	9 600	14 325	28 979	4 595
92	17 564	29 179	44 178	19 742
93	13 182	19 953	47 040	5 148
94	12 228	18 568	31 846	6 469
95	8 635	13 256	31 302	5 376

Correspondance entre le code NAF des Fichiers fonciers et le code NAF de la base Sirene.

Pour le 77, 15 464 locaux possèdent un code Sirene. En parallèle, 10 482 ont exactement le même code Sirene. A l'inverse, 29 332 établissements ne sont pas situés sur une parcelle possédant un code Sirene dans les Fichiers fonciers. Ces chiffres importants au niveau quantitatif ont cependant des explications dues à la structure des bases, à une logique différente, et dans une moindre mesure à des difficultés dans l'appariement des bases.

Connaissance et prévention des risques – Développement des infrastructures – Énergie et climat – Gestion du patrimoine d'infrastructures – Impacts sur la santé – Mobilités et transports – Territoires durables et ressources naturelles – Ville et bâtiments durables

Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement

Direction territoriale Nord-Picardie : 2, rue de Bruxelles - CS 20 275 - 59019 Lille Cedex

Tél : +33 (0)3 20 49 60 00 – fax : +33 (0)3 20 53 15 25

Siège social : Cité des mobilités - 25, avenue François Mitterrand - CS 92 803 - F69674 Bron Cedex - Tél : +33 (0)4 72 14 30 30

Établissement public - Siret 130018310 00016 - TVA Intracommunautaire : FR 94 130018310 www.cerema.fr