

# Enrichissement de l'OCS-GE2 à partir des Fichiers fonciers

Travail sur les emprises d'activité

Avril 2021



Partenaire(s) de l'étude



# Enrichissement de l'OCS-GE2 à partir des Fichiers fonciers

## Travail sur les emprises d'activités

### Historique des versions du document

Version	Date	Commentaire
V1	12/01/21	Version soumise à relecture qualité interne
V2	13/01/21	Version validée suite à relecture qualité interne
V3	15/01/21	Version suite à relecture du CR Grand Est
V4	26/01/21	Version suite à relecture groupe utilisateur
V5	02/02/21	Version finale
V6	06/04/21	Version finale – modification du nom des produits

### Affaire suivie par

<b>Martin Bocquet</b> - Département DATHa – Unité CFU
Tél. : 03 20 49 62 71
Courriel : martin.bocquet@cerema.fr
<b>Site de Lille</b> : Cerema Hauts-de-France – 44 ter rue Jean Bart 59 000 Lille

### Références

n° d'affaire : 20-HF-0081

Partenaire : Conseil Régional Grand Est (contact : Amandine Hamm)

Marché n°2020-00007

Rapport	Nom	Date	Visa
Établi par	Martin Bocquet	12/01/21	
Avec la participation de			
Contrôlé par	Olivier Dupré	13/01/21	
Validé par	Hervé Coulon		

### Résumé de l'étude :

Le Cerema a travaillé à l'enrichissement des emprises d'activité de l'OCS-GE2 de la région Grand Est, en produisant deux bases de données :

- La BD-EA+ est un produit correspondant au redécoupage de l'OCS-GE2 (postes d'emprises d'activités 131) par les Fichiers fonciers. Il s'agit donc de diviser des polygones de type « activité » en des polygones plus précis (commerce, artisanat, etc.).
- La BD-EA est une base de données reprenant la liste des emprises d'activités, ainsi que des informations de caractérisation (surfaces, nombre de locaux, dates de vente, date de première et de dernière construction...). Ces indicateurs pourront être les premiers éléments d'un observatoire régional des emprises d'activités.

Ce présent rapport reprend les modes de création de ces bases, ainsi que les possibilités d'utilisation et les limites de ces bases.

# SOMMAIRE

<b>Principe de l'étude et présentation des livrables.....</b>	<b>4</b>
<b>L'OCS-GE2 de la région Grand Est.....</b>	<b>4</b>
<b>Un besoin de mieux qualifier l'activité économique.....</b>	<b>4</b>
<b>Descriptif des données et contenu du rapport.....</b>	<b>5</b>
<b>Redécoupage des zones 131 (Création d'une BD-EA+).....</b>	<b>6</b>
<b>Descriptif du travail.....</b>	<b>6</b>
<b>Limites et points d'attention.....</b>	<b>7</b>
<b>Définition et fiabilité de la classification.....</b>	<b>7</b>
<b>Étape 1 : affectation de chaque TUP à une classe.....</b>	<b>10</b>
<b>Étape 2 : croisement entre OCS-GE2 et Fichiers fonciers.....</b>	<b>11</b>
<b>Étape 3 : extrapolation.....</b>	<b>12</b>
<b>Conclusion et structuration de la base.....</b>	<b>14</b>
<b>Création de la BD-EA.....</b>	<b>15</b>
<b>Principe de la BD-EA.....</b>	<b>15</b>
<b>Paramètres utilisés.....</b>	<b>15</b>
<b>Mise en place d'indicateurs.....</b>	<b>16</b>
<b>Lien entre l'OCS-GE2, la BD-EA+ et la BD-EA.....</b>	<b>17</b>
<b>Lien entre les données.....</b>	<b>17</b>
<b>Annexe 1: conditions et filtres permettant le classement d'un local.....</b>	<b>18</b>
<b>Annexe 2 : liste et définition des indicateurs de la BD-EA.....</b>	<b>19</b>

# Principe de l'étude et présentation des livrables

---

## L'OCS-GE2 de la région Grand Est

La région Grand Est et l'État, à travers la plate-forme GéoGrandEst, ont lancé en 2019 la production d'une occupation des sols grande échelle sur leur territoire. Cette occupation des sols, l'OCS-GE2, cartographie le territoire selon une nomenclature en 53 postes. L'unité minimale d'interprétation (polygone le plus petit) varie entre 250 m<sup>2</sup> et 1000 m<sup>2</sup> selon les postes.

Cette production, démarrée fin 2019, est réalisée département par département. À ce jour (janvier 2021), le département du Bas-Rhin (67) est diffusé, et les départements de la Moselle (57) et des Vosges (88) le seront prochainement. Les caractéristiques et la documentation technique du produit OCS-GE2 sont présentes sur le site de GeoGrandEst<sup>1</sup>.

---

## Un besoin de mieux qualifier l'activité économique

Pour répondre aux besoins des utilisateurs, l'OCS-GE2 doit permettre de suivre l'occupation du sol liée aux activités économiques. Cependant, l'OCS-GE2 étant réalisée par photo-interprétation, la différenciation entre certains types d'activités n'est pas possible. A titre d'exemple, à partir de photographies aériennes, il est difficilement possible de faire la différence entre un commerce ou des bureaux.

Dans ce contexte, la région a missionné le Cerema en 2020, avec un double objectif :

1 – enrichir l'OCS-GE2, en affinant la nomenclature présente, notamment pour permettre de repérer certaines activités spécifiques ;

2 – utiliser les données produites pour créer une première base de données des emprises d'activités.

Dans les deux cas, il s'agit de retraiter les données en ajoutant des informations issues d'une autre source, les Fichiers fonciers<sup>2</sup>.

## Déroulé de la mission

Cette mission se compose d'une phase préparatoire et d'une phase de production. La phase préparatoire se termine en janvier 2021. Cette phase a eu pour objectif de :

- clarifier les besoins,
- définir le contenu précis des données et leur méthodologie de création
- pointer les avantages, mais aussi les limites de cette donnée
- réaliser le travail d'enrichissement de l'OCS-GE2 pour le département du Bas-Rhin
- constituer la base de données sur les emprises d'activités pour ce département

Pour cela, 3 ateliers avec les utilisateurs ont été organisés, chacun d'entre eux étant précédé par la production d'un jeu de données test.

Le présent rapport clôt la phase préparatoire.

---

<sup>1</sup> [www.geograndest.fr](http://www.geograndest.fr)

<sup>2</sup> Les grandes caractéristiques des Fichiers fonciers, ainsi que les limites de la donnée, seront décrites dans la suite du rapport. Les informations détaillées de cette base sont présentes sur le site Datafoncier: <https://datafoncier.cerema.fr/>

## Descriptif des données et contenu du rapport

Dans ce cadre, le Cerema a produit deux données complémentaires, la BD-EA+ et la BD-EA.

**La BD-EA+** La BD-EA+ est un produit correspondant au redécoupage de l'OCS-GE2 (postes d'emprises d'activités 131) par les Fichiers fonciers. Il s'agit donc de diviser des polygones de type « activité » en des polygones plus précis (commerce, artisanat, etc.).



Polygones marqués comme « activité » dans l'OCS-GE2

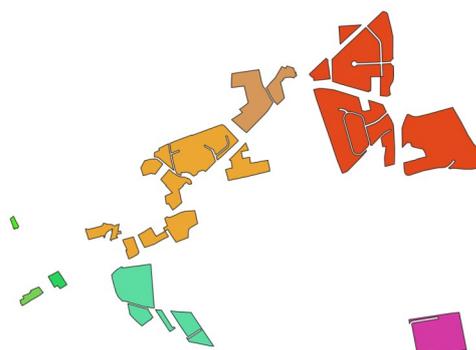


Redécoupage de cette emprise avec les données issues des Fichiers fonciers. Chaque couleur correspond à un type d'activité différent

**La BD-EA** La BD-EA est une base de données reprenant la liste des emprises d'activités, ainsi que des informations de caractérisation (surfaces, nombre de locaux, dates de vente, date de première et de dernière construction...). Ces indicateurs pourront être les premiers éléments d'un observatoire régional des emprises d'activité.



Polygones marqués comme « activité » dans l'OCS-GE2



Regroupement de ces polygones selon les emprises d'activités (1 couleur = 1 emprise d'activité)

**Métadonnées et rapport explicatif** Ces deux bases de données sont accompagnées de métadonnées. Le présent rapport décrit la méthode de création des données, ainsi que les limites liées à celle-ci.

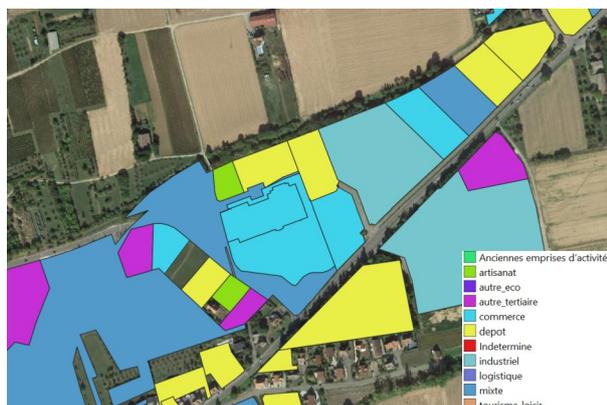
# Redécoupage des zones 131 (Création d'une BD-EA+)

## Descriptif du travail

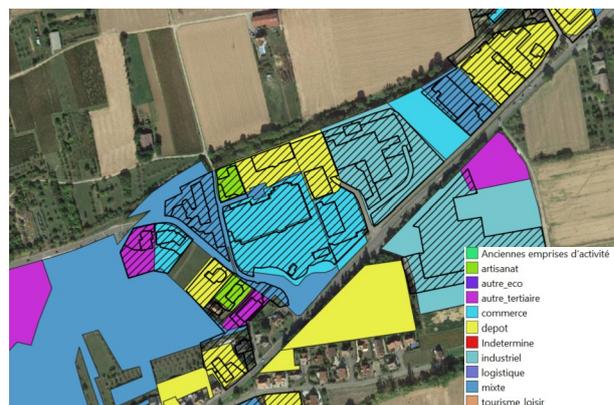
L'objectif de cette partie est d'enrichir l'OCS-GE2, en redécoupant les zones à destination d'activité (zones '131' de l'OCS-GE2) par une typologie plus précise. En d'autres termes, il s'agit de redécouper les polygones « activités » en une multitude de polygones « commerce », « bureaux », « industrie », etc.

Ce redécoupage s'appuie sur les données des Fichiers fonciers (correspondant à l'année N+1 du millésime de l'OCS-GE2), et se déroule en plusieurs étapes :

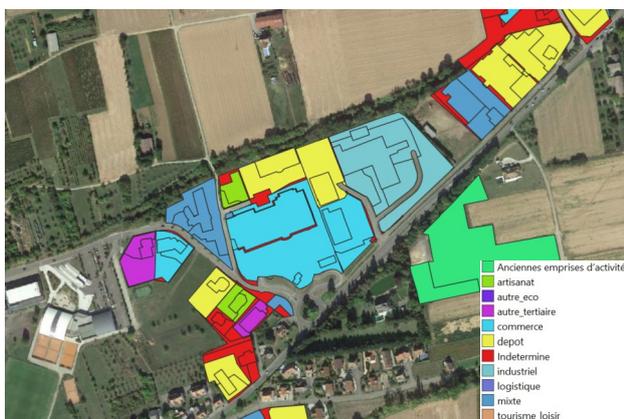
- Étape 1 – les Fichiers fonciers sont retravaillés pour définir une typologie des activités à partir de cette seule base : chaque parcelle se voit assigner un usage majoritaire (artisanat, industrie, commerce...).
- Étape 2 – On découpe l'OCS-GE2 à l'aide de ces parcelles.
- Étape 3 – Une partie des espaces de l'OCS-GE2 ne croise pas des parcelles assignées des Fichiers fonciers. Il s'agit notamment d'espaces « associés » (bords de voirie, parcelles sans locaux tels que les parkings et espaces de stockage...). Dans ce cas, les polygones non affectés sont classés selon les polygones adjacents : un polygone non affecté entouré de commerce sera affecté au commerce. Cette méthode, décrite dans les grandes lignes ci-dessous, est adaptée de travaux précédents, auxquels il est possible de se référer pour connaître la méthodologie détaillée.



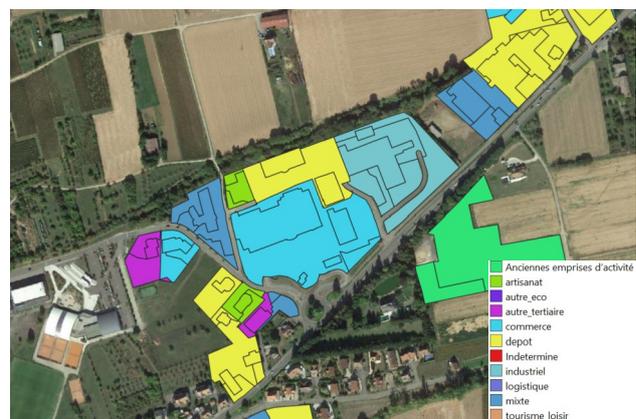
Etape 1 : détermination d'une occupation des sols « activité » selon les Fichiers fonciers



Etape 2a – Croisement de l'OCS-GE2 (en hachuré) avec les Fichiers fonciers



Etape 2b : Découpage de l'OCS-GE2 à l'aide des Fichiers fonciers. Les polygones rouges correspondent au « non affecté » dû aux différences de contour entre les Fichiers fonciers et l'OCS-GE2



Etape 3 – Recalage de la donnée et assignation des parcelles « non affectées » en fonction de la typologie des polygones adjacents.

L'étape 1 est celle qui a le plus d'impact sur l'interprétation des résultats. Elle a donc été détaillée ci-dessous. Les étapes 2 et 3 sont les étapes les plus complexes sur le plan géomatique. Elles sont cependant issues d'une méthodologie antérieure. Dans ce contexte, seules les grandes lignes, ainsi que les points d'attention des étapes 2 et 3 seront décrites. Le lecteur pourra cependant se référer aux travaux antérieurs si besoin<sup>3</sup>.

---

### Limites et points d'attention

Le traitement réalisé est un traitement automatique, sans correction manuelle. Il présente donc les limites suivantes :

- Les données de base, issues des Fichiers fonciers, sont des données fiscales. Ces données sont donc déclaratives, et peuvent présenter des biais liés aux définitions fiscales.
- De la même manière, la donnée de base utilisée est celle de l'OCS-GE2 (emprises 131). L'objectif de cette étude n'était pas de remettre en cause ou de caractériser la donnée de base. Ainsi, la donnée de sortie est aussi tributaire de la donnée OCS-GE2. En particulier, les espaces non présentés comme à usage d'activité dans l'OCS-GE2 n'ont pas été caractérisés. De même, les limites globales des zones n'ont pas été modifiées. En conséquence, les équipements collectifs au sens de l'OCS-GE2 (codes 121) n'ont pas été redécoupés.
- Au niveau du traitement géomatique, le traitement redécoupe les zones OCS-GE2 par les données des Fichiers fonciers. Cela peut donc, in fine, présenter des découpages parfois problématiques. Des méthodes de retraitement ont été utilisées pour limiter ces artefacts, mais certains peuvent subsister.

---

### Définition et fiabilité de la classification

Les Fichiers fonciers sont une base de données issues des données fiscales liées à la taxe foncière. On dispose de ces données à l'échelle des locaux, des parcelles et des propriétaires. L'objectif de l'étude est ici d'assigner, à chacune des parcelles d'activité, une typologie précise.

#### Un travail à l'unité foncière

Une des limites des Fichiers fonciers réside dans leur prise en compte des unités foncières. En effet, lorsque plusieurs parcelles de même propriétaire sont adjacentes, la totalité des informations est souvent assignée à une parcelle. Ainsi, si deux parcelles adjacentes de même propriétaire ont chacune une maison, il est possible qu'une parcelle soit affichée comme contenant 2 maisons, et que l'autre n'en possède aucune. Cela pose problème, dans le cadre de l'activité, pour réussir à définir une typologie.

Pour fiabiliser la base, il est donc nécessaire de travailler à l'échelle de l'unité foncière. En d'autres termes, les parcelles seront regroupées à l'unité foncière, qui constituera in fine la plus petite entité fiable. L'unité de base utilisée sera donc la « TUP » (regroupement de toutes les parcelles en unité foncière).

---

3 Les méthodes sont décrites dans les rapports suivants : <https://datafoncier.cerema.fr/usages/consommation-des-espaces-et-occupation-des-sols/enrichissement-ocsge-par-les-fichiers-fonciers>  
<https://datafoncier.cerema.fr/usages/consommation-des-espaces-et-occupation-des-sols/enrichissement-ocsge-ign-fichiers-fonciers-methode-appfondie-sur-grand-territoire>

### **Définition de la classification**

Chaque entité issue des Fichiers fonciers sera classée selon une des 11 catégories dans le tableau ci-dessous. Parmi celles-ci se trouve une catégorie « indéterminée », qui regroupe les polygones classés comme « activité » au sens de l'OCS-GE2, mais ne recoupant pas de TUP classée comme « activité ». Les cas pour lesquels les polygones sont finalement classés en « indéterminés » sont listés dans l'étape 3.

La classification s'appuie donc en grande partie sur les Fichiers fonciers, et plus précisément sur les informations suivantes contenues dans la base, pour chaque local professionnel :

- Le type d'activité du local. Ce type d'activité, se sépare en 38 catégories, allant du magasin au dépôt en passant par les locaux touristiques<sup>4</sup>. Ce type d'activité concerne la forme du local, c'est-à-dire ce pourquoi le local a été créé. Ce champ est fiable, mais la typologie n'est cependant pas toujours adaptée aux besoins.
- Le code NAF de l'occupant du local. Ce code, plus précis, n'est cependant pas toujours rempli. Pour cette raison, on privilégie la variable précédente si cela est possible.
- Certains locaux industriels, souvent d'une grande ampleur, disposent d'un mode de taxation différent. Sans rentrer dans les détails<sup>5</sup>, ces locaux sont bien répertoriés dans les Fichiers fonciers, mais les informations liées à ces locaux ne sont pas toujours fiables. Ainsi, le local sera bien répertorié en tant que « industrie », mais sa surface sera mal remplie (souvent égale à 0 m<sup>2</sup>).

Ainsi, la méthodologie mobilise à la fois le type d'activité, le code NAF, et cette information liée aux locaux industriels.

### **Définition de la fiabilité**

On évalue de plus la fiabilité de chacune des catégories produites. Attention, cette fiabilité est aujourd'hui déterminée a priori par le Cerema, avant toute remontée réalisée par les territoires. Cette fiabilité est donc réalisée « à dire d'expert », en s'appuyant sur les connaissances acquises sur les Fichiers fonciers. Dans ce contexte, les déclarations doivent être prises avec du recul, et seront amenées à évoluer en fonction des retours.

De la même manière, ce classement est relatif : une fiabilité de 2 / 5 signifie que le classement est moins fiable qu'un 3 / 5, et non que le classement n'est pas fiable dans l'absolu.

La surface indiquée correspond à la surface de la donnée finale, après réalisation de la phase 3.

---

4 La typologie précise, ainsi que les correspondances entre chaque catégorie, est développée dans le formulaire de déclaration de la taxe foncière, dont la notice est présente à cette adresse :

[https://www.impots.gouv.fr/portail/files/formulaires/6660-rev/2019/6660-rev\\_1856.pdf](https://www.impots.gouv.fr/portail/files/formulaires/6660-rev/2019/6660-rev_1856.pdf)

5 Pour plus d'information, il est possible de se référer au document suivant : <https://datafoncier.cerema.fr/les-champs-lies-la-taxation-fonciere-dans-les-fichiers-fonciers>

Catégorie	% de la surface des emprises (dep 67)	Fiabilité	Commentaire
artisanat	10,66 %		Cette catégorie reprend toutes les parcelles à usage d'artisanat (garages, menuiseries...). Certaines classifications peuvent cependant être soumises à interprétation. A titre d'exemple, un garage automobile pourrait, selon son activité, être classé dans la catégorie « commerce » ou en « artisanat ». De la même manière, la frontière entre artisanat et industrie peut être assez poreuse.
autre_eco	0,03 %		Cette catégorie reprend les locaux particuliers au sens des impôts, c'est-à-dire ne rentrant dans aucune des autres catégories. Ces polygones sont bien couverts par les Fichiers fonciers, mais les informations des Fichiers fonciers les classent dans une catégorie « divers ».
bureau_et_tertiaire	8,03 %		Cette catégorie reprend les bureaux, les établissements d'enseignement et les établissements hospitaliers. Il faut noter que cette catégorie est sans doute sous-estimée. En effet, de nombreux bureaux sont situés dans le tissu urbain classique, et peuvent être confondus avec des logements. De la même manière, l'OCS-GE2 reprend, dans sa typologie, des nomenclatures spécifiques pour les emprises universitaires (1211) et hospitalières (1212).
commerce	14,37 %		Cette catégorie reprend toutes les parcelles à usage de commerce. Il s'agit donc de tous les locaux dont l'activité principale est la vente. Cette catégorie est globalement fiable. Encore une fois, certains cas très particuliers peuvent être observés, notamment sur des surfaces reprenant une surface de vente et une surface de bureau. Dans ces cas particuliers, la parcelle est très souvent classée en « commerce ». Il peut toutefois subsister certains cas pour lesquels ces espaces sont classés en « bureaux ». Enfin, il faut rappeler que de nombreux commerces, notamment en centre-ville, sont situés dans des immeubles mixtes, et donc non repérés en tant que emprise d'activité de l'OCS-GE2.
depot	10,89 %		Les dépôts sont constitués des parkings et espaces de stockage, lorsque ceux-ci sont dissociables des autres locaux. En d'autres termes, si une parcelle contient du commerce et un parking, celle-ci sera classée en « commerce ». S'il s'agit de deux parcelles différentes, l'une sera classée en commerce, l'autre en dépôt.
Anciennes emprises d'activités	6,05 %		La catégorie reprend les données de l'OCS-GE2, (catégorie « 1314 »).
Indetermine	4,40 %		Cette catégorie reprend les polygones pour lesquels aucune assignation n'a pu être définie. Il s'agit donc d'une catégorie par défaut.
industriel	31,53 %		L'industrie reprend les parcelles industrielles. Au niveau des données, est considéré comme industriel tout ce qui nécessite un outillage important. Dans ce cadre, la frontière entre catégories artisanat et industriel est parfois assez poreuse.
logistique	5,37 %		La logistique est définie à partir des codes NAF présents dans les Fichiers fonciers. Cette donnée est fiable lorsqu'elle est présente. Cependant, le code NAF n'est pas toujours renseigné. Cette catégorie pourrait donc ne pas être exhaustive.
mixte	7,25 %		Cette catégorie s'applique lorsque deux locaux de deux typologies différentes existent sur la même parcelle, mais qu'aucun n'est prépondérant. Pour plus d'information, il est conseillé de se référer au paragraphe « passage de la typologie à la TUP » décrit à l'étape 1.
tourisme_loisir	1,42 %		Cette catégorie reprend les établissements de spectacles, salles de sport, théâtres, hôtels, campings, locaux culturels ou locaux à usage de culte. La fiabilité est plutôt bonne. Il est à noter que l'OCS-GE2 dispose d'une catégorie reprenant les équipements sportifs et de loisirs (code 1213).

Ces éléments de fiabilité sont présentés pour permettre une meilleure appropriation des données. Il s'agit d'une évaluation a priori, qui sera revue après un certain recul sur l'utilisation de la donnée. Le classement est relatif : un classement en « orange » ne veut pas dire que la donnée est inutilisable, mais seulement qu'elle sera moins fiable que les autres catégories. Enfin, ce classement se veut le plus prudent possible. A titre d'exemple, aucune catégorie n'a bénéficié de la fiabilité maximale. Après un certain recul sur les données, certaines catégories pourraient être réévaluées pour une meilleure fiabilité.

#### Commentaire

	Ces données sont très fiables, exhaustives, et ne nécessitent pas d'être confrontées à des données terrains.
	Ces données sont par construction très fiables, mais peuvent parfois présenter des décalages par rapport à une confrontation terrain. Ces décalages peuvent être considérés comme non significatifs à une échelle communale.
	Ces données sont fiables et utilisables mais peuvent ne pas être exhaustives ou bien présenter des limites d'utilisation.
	Ces données peuvent présenter des soucis de classement, leur modalité de détermination en faisant souvent des catégories par défaut.
	Ces données doivent être considérées comme indicatives et leurs fiabilités comme incertaines. Les variables ne peuvent être à nouveau croisées avec d'autres données.

---

## Étape 1 : affectation de chaque TUP à une classe

Cette étape se décompose elle-même en deux temps :

- d'abord l'affectation de chaque local à une classe
- ensuite, le passage de « l'échelle du local » à « l'échelle TUP »

### *Affectation de chaque local à une classe*

Chaque local est classé selon la typologie décrite dans cette étape 1. Ce classement s'appuie sur les 3 éléments présents dans les Fichiers fonciers et décrits ci-dessus, à savoir :

- le code catégorie du local (champ « typeact »),
- le code NAF de l'occupant du local (champ « cconac »),
- les données liées au type de taxation réalisé (champ « ccoeva »).

En sus, pour les données des friches, nous utilisons les données de base de l'OCS-GE2 (poste anciennes emprises d'activité 1314). Les conditions techniques (champs et filtres utilisés) sont présentées dans l'annexe 1.

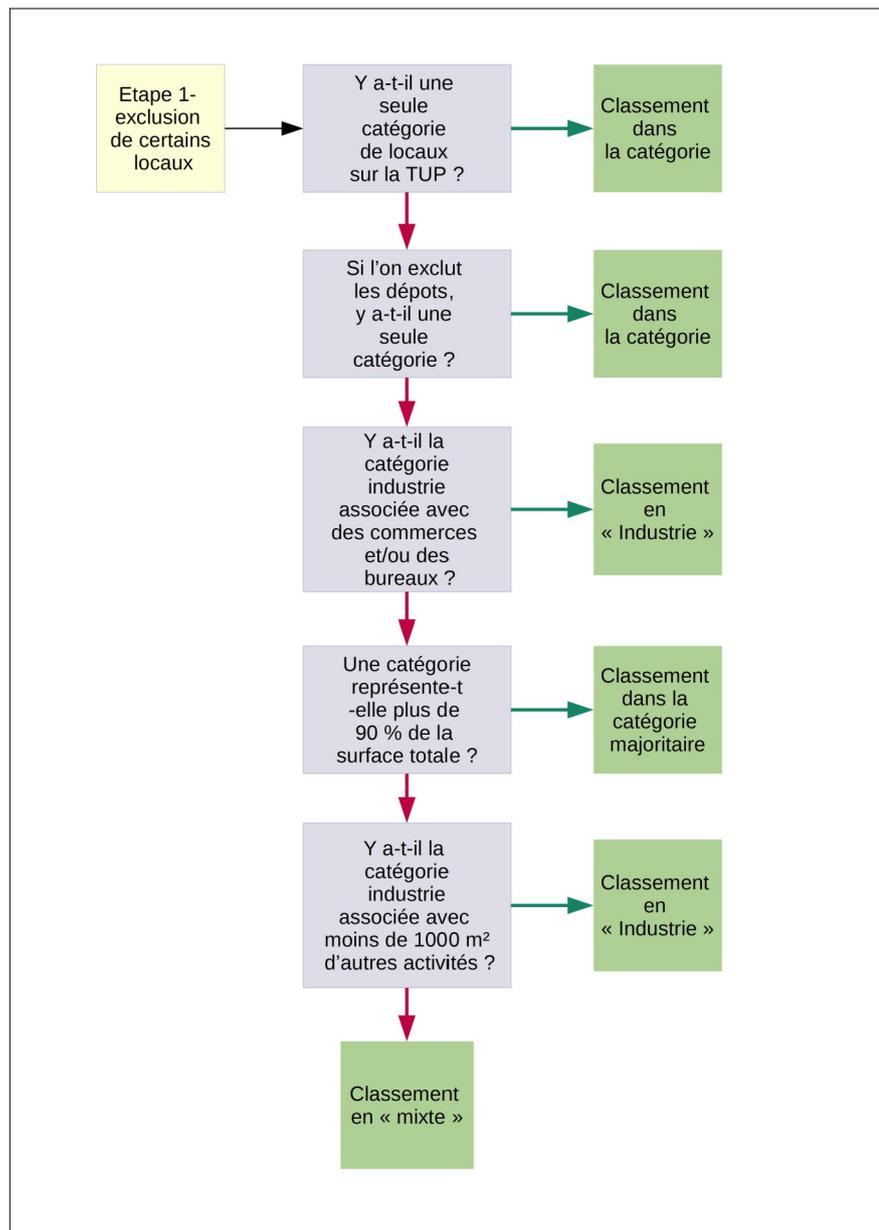
### *Passage de la typologie à la TUP*

Une fois les locaux classés, l'objectif est de transformer des données au local en des données à la TUP. A titre d'exemple, il peut y avoir, sur une même TUP, 5 locaux de bureaux et un local de commerce. L'objectif est de voir si, même avec des locaux différents, la TUP serait classée ou non en mixte.

En effet, si une TUP contient 2500 m<sup>2</sup> de surface de commerce, et 80 m<sup>2</sup> de bureaux, il serait logique que cette TUP soit classée en « commerce » et non en « mixte ».

Pour passer d'un classement au local à un classement à la TUP, on réalise, dans l'ordre, le traitement suivant. Celui-ci est résumé dans le logigramme ci-dessous.

- Exclusion de certains locaux du calcul (antennes téléphoniques, chantiers et transformateurs électriques).
- Si une TUP ne possède que des locaux appartenant à une catégorie, cette TUP est classée dans cette catégorie. Ainsi, si une TUP possède 3 locaux classés en « commerce », la TUP complète sera classée en « commerce ».
- On considère en outre que la catégorie « depot » n'est pas prioritaire. Ainsi, si la TUP possède un classement « artisanat » et un classement « depot », elle sera classée en « artisanat ». En effet, la catégorie « depot » est majoritairement constituée de parkings et d'espaces de stockage, qui ne sont donc pas constitutifs d'une typologie en particulier.
- Pour les TUP possédant de l'industrie et des bureaux (ou de la logistique et des bureaux), on considère qu'elle sera à usage d'industrie (ou de logistique). En effet, les industries sont évaluées différemment des autres locaux dans les Fichiers fonciers, et les surfaces liées sont souvent peu fiables. Cependant, ce mode d'évaluation n'est mobilisé que pour des industries d'une certaine taille. Dans ce contexte, il est donc préférable de postuler que les TUP de cette sorte (mixité industrie/bureaux ou industrie/logistique) soient à usage dominant d'industrie.
- Ensuite, si une typologie représente plus de 90 % de la surface bâtie totale (ex : 950 m<sup>2</sup> de bureaux et 50 m<sup>2</sup> de commerce), on classe la TUP dans cette catégorie.
- Enfin, les TUP industrielles possédant moins de 1000 m<sup>2</sup> d'autres catégories sera classée en industrielle.
- Si après ces différents filtres, la TUP n'a pas été assignée, elle sera classée en « mixte ».



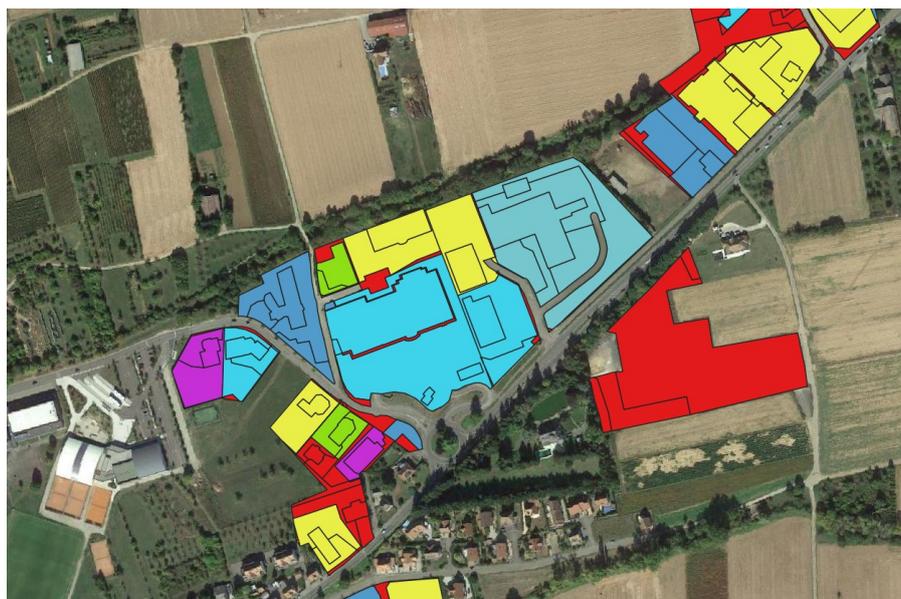
Logigramme détaillant le passage d'une typologie de locaux vers une typologie à la TUP

## Étape 2 : croisement entre OCS-GE2 et Fichiers fonciers

Cette étape consiste en un simple croisement géomatique entre les deux entités. En d'autres termes, les polygones résultants seront le plus petit dénominateur commun entre la partie « activité » de l'OCS-GE2 (c'est-à-dire les polygones « 131 »). Sur les zones en commun entre l'OCS-GE2 et les Fichiers fonciers, on assigne la typologie définie précédemment.

Ce croisement appelle cependant plusieurs remarques :

- La couche de base utilisée est bien celle de l'OCS-GE2 de niveau 5 (comprenant le degré d'imperméabilisation des sols). Dans ce cadre, aucun changement n'a été réalisé sur le contour.
- L'OCS-GE2 et les Fichiers fonciers n'ont pas les mêmes contours. Dans ce cadre, on observe donc des polygones de l'OCS-GE2 qui ne sont pas assignés par les Fichiers fonciers.
- Le croisement préserve les différences originelles, sur l'OCS-GE2, entre les zones imperméables, perméables et bâties.



**Exemple de résultats après croisement des couches. Les polygones rouges correspondent aux couches de l'OCS-GE2 non couvertes par les Fichiers fonciers**

### Étape 3 : extrapolation

L'objectif de cette étape est de permettre de transformer le maximum de polygones indéterminés en des polygones affectés.

La méthodologie générale est issue de la méthode développée par le Cerema<sup>6</sup>. Elle consiste à assigner, pour chaque polygone, l'affectation majoritaire des polygones adjacents : si un polygone non affecté a 20 m de linéaire commun avec un polygone « commerce » et 200 m avec un polygone « depot », ce polygone sera affecté en « depot ». Il est à noter que l'assignation ne travaille que sur les polygones directement adjacents (moins de 0,5 m). Ainsi, si un polygone est situé de l'autre côté d'une route, il ne sera pas concerné.

Des traitements sont en outre réalisés pour supprimer les polygones trop petits (en dessous de l'Unité Minimale d'Interprétation), redécouper certains polygones indéterminés trop importants, et corriger les erreurs géométriques (géométries illégales, problèmes de topologie...). Pour connaître le détail de ces traitements complexes, il est possible de se référer aux rapports cités ci-dessus.



**Couche issue du croisement Fichiers Fonciers / OCS-GE2**



**Résultat suite à l'assignation**

<sup>6</sup> Les méthodes sont décrites dans les rapports suivants : <https://datafoncier.cerema.fr/usages/consommation-des-espaces-et-occupation-des-sols/enrichissement-ocsge-par-les-fichiers-fonciers>  
<https://datafoncier.cerema.fr/usages/consommation-des-espaces-et-occupation-des-sols/enrichissement-ocsge-ign-fichiers-fonciers-methode-appfondie-sur-grand-territoire>

**À quoi correspondent les polygones « indéterminés » restants ?**

Il existe des espaces classés en tant qu'espaces d'activités, mais ne contenant pas de locaux dans les Fichiers fonciers. Les surfaces « indéterminées » correspondent à 4,4 % des surfaces d'activités sur le département 67.

Les polygones classés comme « indéterminés » peuvent représenter :

- 1 - les bâtiments agricoles, qui peuvent parfois être classés en tant qu'espaces d'activités. Les Fichiers fonciers ne contenant pas de données liées aux activités agricoles, les bâtiments agricoles ne pourront être enrichis.
- 2 - Sur certains bâtiments récents, il peut y avoir un décalage entre la construction effective (donc visible sur les photographies aériennes) et la déclaration fiscale des Fichiers fonciers (local encore non déclaré).
- 3 - Des parkings et espaces de stockage isolés. Dans certaines emprises d'activités, les parkings peuvent être situés de l'autre côté de la route. Dans ce cas, le traitement ne prenant en compte que les polygones adjacents, l'assignation ne peut se faire.
- 4 - Un croisement problématique entre la géométrie de l'OCS-GE2 et les Fichiers fonciers.
- 5 - Une différence entre les définitions fiscales et l'interprétation du photo-interprète. A titre d'exemple, on observe des maisons individuelles qui réalisent des activités de stockage. À cause de ce stockage, le photo-interprète reconnaît cet espace comme à usage d'activité.



**Cas 1 : bâtiment agricole**



**Cas 2 : local neuf**



**Cas 3 : parking / stockage**



**Cas 4 : pas de lien entre FF et OCSGE**



Cas 5 : Maison avec une aire de stockage

### Conclusion et structuration de la base

La BD-EA+ est donc une base de données :

- reprenant les contours des emprises d'activité de l'OCS-GE2 (code 131) ;
- découpant ceux-ci selon une typologie en 11 postes (commerce, artisanat...);
- gardant la différence entre les espaces imperméabilisés, non imperméabilisés et bâtis.

La BD-EA+ est composée des 5 champs suivants :

Nom du champ	Type	Commentaire
fid	integer	Identifiant BD-EA+ du polygone
gid	Integer	Champ reprenant les identifiants des polygones de l'OCS-GE2. À un polygone de l'OCS-GE2 correspond un ou plusieurs polygones de la BD-EA+ (voir la section consacrée pour plus d'information).
l_idcom	Texte[] (tableau de texte)	Ce champ reprend les codes INSEE des communes. Dans la majeure partie des cas (> 99%), chaque polygone appartient à une seule commune. Il peut y avoir des polygones à cheval sur plusieurs communes.
Typologie	Texte	Typologie en 11 postes définie précédemment
cod_n5	Numeric	Ancienne typologie issue de l'OCS-GE2. Ce champ peut être réutilisé, en reprenant le dernier chiffre, pour faire la différence entre imperméable bâti (131X1), imperméable non bâti (131X2) et perméable (131X3).

Cette donnée peut donc être utilisée, en tant que telle, comme socle pour les études locales, en complétant la connaissance des emprises d'activités de l'OCS-GE2 de niveau 4. En outre, elle sert de socle pour la construction d'une première base de données d'observation des emprises d'activités à une échelle départementale, la BD-EA.

# Création de la BD-EA

## Principe de la BD-EA

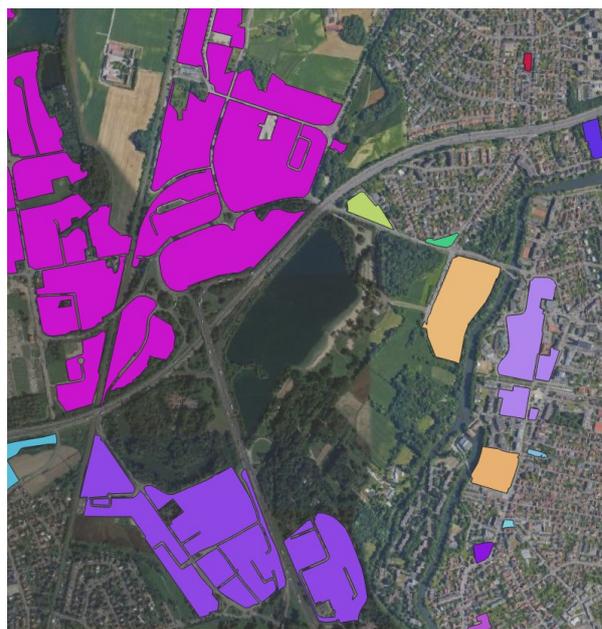
La méthode de création de la BD-EA+ permet d'observer, à une échelle fine, l'occupation des sols liée au type d'activité.

Cependant, ces polygones ne correspondent pas à une emprise d'activité au sens urbanistique du terme. Pour se rapprocher de celle-ci, il est nécessaire de regrouper les polygones adjacents.

Pour cette base de données, le terme « emprise d'activité » a été privilégié au terme « zone d'activité ». La BD-AE reprend les contours des emprises d'activité de la BD-EA+ (et de l'OCS-GE2).



Donnée de base (OCS-GE2) sur la partie « activité »



Repérage des emprises d'activités (1 couleur = 1 emprise) par union de polygones proches.

## Paramètres utilisés

Pour créer une emprise d'activité, nous avons fusionné les polygones situés à moins de 20 m les uns des autres. Un post-traitement a ensuite été utilisé afin de supprimer les « trous » dans les polygones.

### Choix de la distance de 20 m

Lors de la phase d'étude, s'est posée la question de la distance à retenir pour lier les polygones. En d'autres termes, si deux espaces d'activités sont situés à moins de X m les uns des autres, constituent-ils une « emprise d'activité », au sens urbanistique ou en termes de gestion ?

Pour choisir ce paramètre, diverses simulations, avec différentes distances (2m, 10 m, 20 m et 50 m) ont été proposées aux utilisateurs dans le cadre d'ateliers. En outre, d'autres tests ont été réalisés, notamment avec une distance proportionnelle à la taille de la commune, ou proportionnelle à la taille des polygones d'origine.

L'objectif est d'utiliser une distance suffisamment grande pour regrouper ensemble les différents éléments d'une grande emprise d'activité, mais de ne pas regrouper les petites entités de centre-ville.

Suite aux retours des utilisateurs, il s'est avéré que la distance de 20 m constituait le meilleur compromis.

---

## Mise en place d'indicateurs

Une fois le principe d'emprise d'activité déterminé, chaque polygone est enrichi par une série d'indicateurs, issus de la BD-EA+ et des Fichiers fonciers.

Les indicateurs sont de plusieurs types :

- surface géographique, c'est-à-dire surface des polygones inscrits dans la BD-EA+ ;
- surface des locaux ;
- surfaces imperméabilisées, imperméabilisées bâties et non imperméabilisées ;
- date de première et de dernière construction dans l'emprise d'activité ;
- nombre de changements de propriétaires ;
- liste de codes NAF des activités existantes.

La liste exacte des indicateurs, ainsi que leur définition, est présente dans l'annexe 2.

Une partie de ces indicateurs est déclinée en fonction de la typologie précédente. A titre d'exemple, nous avons la date de première construction de l'emprise d'activité, mais aussi la date de première construction d'un commerce ou d'un local touristique sur la zone.

### *Limites de l'approche*

Les indicateurs inscrits dans la base sont issus de la BD-EA+, et souffrent donc des mêmes limites que celle-ci. De plus, d'autres indicateurs sont issus des Fichiers fonciers. Pour ceux-ci, plusieurs limites existent :

1- le croisement entre la BD-EA et les Fichiers fonciers se fait sur une base géographique. En d'autres termes, la BD-EA reprend l'intégralité des informations de la parcelle, et ce même si ce croisement n'est que partiel. De manière très rare, il serait possible que les informations d'une TUP soient reprises par 2 emprises d'activités.

2 – Les informations des Fichiers fonciers liées à l'industrie peuvent manquer de fiabilité. En particulier, les surfaces des locaux et les dates de constructions, sur la partie « activité », sont peu fiables.

## Lien entre l'OCS-GE2, la BD-EA+ et la BD-EA

### Lien entre les données

Les 3 bases de données ont chacune un identifiant unique pour chaque polygone :

- gid pour l'OCS-GE2,
- fid pour la BD-EA+,
- id\_ea pour la BD-EA.

Les liaisons entre ces bases sont de type 1-N : à un élément correspond entre 1 et N éléments d'une autre base

### Lien 1-N entre l'OCS-GE2 et la BD-EA+

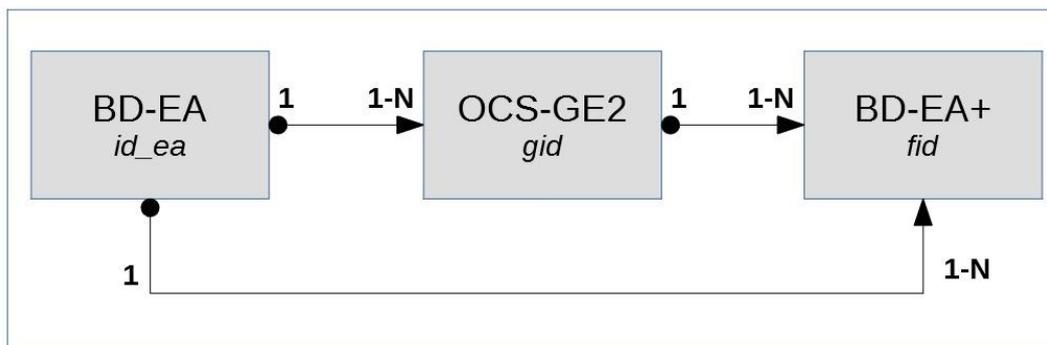
Le process de création de la BD-EA+ se fait en découpant les polygones de l'OCS-GE2. À un polygone de l'OCS-GE2 correspond un ou plusieurs identifiants de la BD-EA+.

La BD-EA+ comprend un champ gid, qui sert de clé étrangère.

### Lien 1-N entre BD-EA et OCS-GE2

La BD-EA reprend un ou plusieurs éléments de l'OCS-GE2, et donc un ou plusieurs éléments de la BD-EA+. Chaque id\_ea contiendra entre 1 et N polygones de l'OCS-GE2.

La BD-EA contient le champ gid, qui sert d'élément de liaison avec l'OCS-GE2.



Modèle Conceptuel de Données (MCD) simplifié

## Annexe 1: conditions et filtres permettant le classement d'un local

Si	Alors	Commentaire
Type d'activité = 'MAG_'	Classement en « commerce »	Cette catégorie reprend toutes les activités commerciales, allant des boutiques en centre-ville aux centres commerciaux, hypermarchés, stations services et halles de marché.
Type d'activité = 'SPE_' ou 'HOT_'	Classement en « tourisme_loisir »	Cette catégorie reprend les établissements de spectacles, salles de sport, théâtres, hôtels, campings, locaux culturels ou locaux à usage de culte.
Type d'activité = 'ATE1' ou 'ATE2'	Classement en « artisanat »	Cette catégorie reprend les locaux exploités à des artisans ou des activités de transformation. A titre d'exemple, les garages automobiles, menuiseries artisanales se situent dans cette catégorie.
Type d'activité = 'DEP_'	Classement en « depot »	La catégorie « depot » reprend les lieux de stockage et entrepôts sans activité logistique. Elle reprend en outre les parkings.
Type d'activité = 'ATE3' ou 'BUR_' ou 'ENS_' ou 'CLI_'	Classement en « bureau_et_tertiaire »	Cette catégorie reprend les bureaux ordinaires, les activités des professions libérales, ainsi qu'une partie des activités médicales (cabinets de médecin, laboratoires d'analyse médicale...). Les cliniques et établissements d'enseignement sont situés dans cette catégorie.
Type d'activité like 'IND_' ou imposition liée à une activité industrielle	Classement en « industriel »	Cette catégorie reprend les locaux industriels, c'est-à-dire usines ou locaux de transformation. La frontière avec la catégorie « artisanat » n'est pas toujours très claire, la différenciation étant normalement réalisée sur la quantité d'outillage nécessaire. Les opérations d'extraction (notamment carrières) sont classées dans cette catégorie.
Code NAF relatif à la logistique <sup>7</sup>	Classement en « logistique »	Cette catégorie reprend les locaux liés à la logistique ou au transport de voyageurs ou de marchandise.
Type d'activité = 'EXC1'	Classement en « autre_eco »	Cette catégorie correspond à des locaux ne pouvant pas être classés dans une catégorie spécifique.
Classement en friche dans l'OCS-GE <sup>8</sup>	Classement en « ancienne_emprise_d'activitef »	Emprises originales issues de l'OCS GE2 (poste 1314)

<sup>7</sup> C'est-à-dire locaux dont le code NAF commence par 49, 51, 52 ou 53.

<sup>8</sup> C'est-à-dire cod\_n4 = 1314

## Annexe 2 : liste et définition des indicateurs de la BD-EA

Nom du champ	Type	Commentaire	Source	Fiabilité
id_zs	integer	Identifiant de la zone d'activité		
l_gid	text	liste des identifiants des polygones de l'OCS-GE2, séparés par une virgule	OCS-GE2 base	
l_insee	text	liste des identifiants INSEE des communes, séparés par une virgule (ex : « 67001,67002 »)		
surf_geom	numeric	Surface géométrique de la zone d'activité	OCS-GE2 base	
area_artisanat	numeric	Surface géométrique de la zone d'activité dédiée à l'artisanat (voir la méthodologie)	OCS-GE2+	3 / 5
area_autre_eco	numeric	idem	OCS-GE2+	2 / 5
area_autre_tertiaire	numeric		OCS-GE2+	3 / 5
area_commerce	numeric		OCS-GE2+	4 / 5
area_depot	numeric		OCS-GE2+	2 / 5
area_indetermine	numeric		OCS-GE2+	1 / 5
area_industriel	numeric		OCS-GE2+	4 / 5
area_mixte	numeric		OCS-GE2+	3 / 5
area_tourisme_loisir	numeric		OCS-GE2+	4 / 5
area_logistique	numeric		OCS-GE2+	2 / 5
area_friche	numeric		OCS-GE2 base	
area_imperbati	numeric	Surface imperméable bâtie (surface zones 131X1 de l'OCSGE2)	OCS-GE2 base	
area_impnonb	numeric	Surface imperméable non bâtie (surface zones 131X2 de l'OCSGE2)	OCS-GE2 base	
area_ninimper	numeric	Surface non imperméabilisée (surface zones 131X3 de l'OCSGE2)	OCS-GE2 base	
nblocal	numeric	nombre de locaux selon les Fichiers fonciers sur la ZA	Fichiers fonciers	4 / 5
jannatmin	integer	année de première construction de la zone (0 = non connu, -1 = non bâti)	Fichiers fonciers	3 / 5
jannatmax	bigint	année de dernière construction de la zone	Fichiers fonciers	3 / 5
dcntpa	numeric	surface des parcelles cadastrées dans la zone d'activité	Fichiers fonciers	5 / 5
dcntarti	numeric	Surface artificialisée cadastrée dans la zone d'activité	Fichiers fonciers	4 / 5
dcntnaf	double precision	Surface NAF cadastrée de la zone d'activité	Fichiers fonciers	4 / 5
sprincp	numeric	Surface principale bâtie des locaux d'activité	Fichiers fonciers	4 / 5
ssecp	numeric	Surface secondaire couverte des locaux d'activité	Fichiers fonciers	4 / 5
ssecncp	numeric	Surface secondaire non couverte des locaux d'activité	Fichiers fonciers	4 / 5
sparkp	numeric	Surface des parkings couverts des locaux d'activité	Fichiers fonciers	4 / 5
sparkncp	numeric	Surface des parkings non couverts des locaux d'activité	Fichiers fonciers	4 / 5
slocal	numeric	Surface totale de tous les locaux	Fichiers fonciers	4 / 5
moyenne_surface_tup	numeric	Surface moyenne des TUP sur la zone d'activité	Fichiers fonciers	3 / 5
densite_batie	numeric	Densité bâtie (surface totale des locaux divisée par la surface des TUP)	Fichiers fonciers	3 / 5
slocal_artisanat	numeric	surface totale des locaux dédiés à l'artisanat	Fichiers fonciers	3 / 5
slocal_autre_eco	numeric	idem		2 / 5
slocal_autre_tertiaire	numeric			3 / 5
slocal_commerce	numeric			4 / 5
slocal_depot	numeric			2 / 5
slocal_indetermine	numeric			1 / 5
slocal_industriel	numeric			4 / 5
slocal_mixte	numeric			3 / 5
slocal_tourisme_loisir	numeric			4 / 5
slocal_logistique	numeric			2 / 5
datmin_artisanat	integer	Date de première construction des locaux dédiés à l'artisanat (0 = non connu, -1 = non bâti)	Fichiers fonciers	3 / 5
datmin_autre_eco	integer	idem		2 / 5
datmin_autre_tertiaire	integer			3 / 5
datmin_commerce	integer			3 / 5
datmin_depot	integer			2 / 5
datmin_indetermine	integer			1 / 5
datmin_industriel	integer			3 / 5
datmin_mixte	integer			3 / 5
datmin_tourisme_loisir	integer			3 / 5
datmin_logistique	integer			2 / 5
datmax_artisanat	bigint	Date de dernière construction des locaux dédiés à l'artisanat (0 = non connu, -1 = non bâti)	Fichiers fonciers	3 / 5
datmax_autre_eco	bigint	idem		2 / 5
datmax_autre_tertiaire	bigint			3 / 5
datmax_commerce	bigint			3 / 5
datmax_depot	bigint			2 / 5
datmax_indetermine	bigint			1 / 5
datmax_industriel	bigint			3 / 5
datmax_mixte	bigint			3 / 5
datmax_tourisme_loisir	bigint			3 / 5
datmax_logistique	bigint			2 / 5
mutation_avant_2000	numeric	Nombre de locaux ayant été vendus pour la dernière fois avant 2000	Fichiers fonciers	4 / 5
mutation_apres_2001	numeric	Nombre de locaux ayant été vendus pour la dernière fois entre 2001 et 2010	Fichiers fonciers	4 / 5
mutation_apres_2011	numeric	Nombre de locaux ayant été vendus pour la dernière fois entre après 2010	Fichiers fonciers	4 / 5
l_typologie	text[]	Liste des typologies d'activité présentes sur la zone		
l_code_naf	character varying[]	Liste des codes NAF des activités présentes sur la ZA	Fichiers fonciers	3 / 5
nb_prop	bigint	nombre de propriétaires	Fichiers fonciers	5 / 5
geom	geometry	géométrie de la zone d'activité	Fichiers fonciers	





**Cerema Hauts-de-France**

44 Ter , rue Jean BART - CS 20275 - 59019 LILLE Cedex  
Tél : 03 20 49 60 00 – Fax : 03 20 53 15 25 – Courriel : [DTerNP@cerema.fr](mailto:DTerNP@cerema.fr)

[www.cerema.fr](http://www.cerema.fr)