



Constitution d'une base de données économiques 2010.

Méthodologie – v2

– Juillet 2013 –

Sommaire

1. IDENTIFICATION DU DOCUMENT	3
1.1. INFORMATIONS GENERALES.....	3
1.2. HISTORIQUE DES EVOLUTIONS.....	3
1.3. CIRCUIT DE VALIDATION.....	3
2. INTRODUCTION.....	4
3. ZONES D'ACTIVITES INDUSTRIELLES ET COMMERCIALES	4
3.1. DEFINITION.....	4
3.2. GEOMETRIE	5
3.2.1. Données existantes.....	5
3.2.2. Données sources supplémentaires	6
3.2.2.1. Occsol 2010	6
3.2.2.2. Données Navteq	6
3.2.3. Cas des polygones délimités par plusieurs types de données.....	7
3.2.3.1. Données Navteq et Occsol	7
3.2.3.2. Navteq, Occsol et données existantes	8
3.3. STRUCTURE DE LA TABLE	9
3.4. RENSEIGNEMENT DES ATTRIBUTS.....	9
3.4.1. Identifiants du site.....	9
3.4.2. Identifiant du pôle	10
3.4.3. Nom du site	10
3.4.4. Surface.....	10
3.4.5. Saisie.....	10
3.4.6. Référence spatiale et précision géométrique	10
3.4.7. Nombre d'entreprises.....	10
3.4.8. Effectifs.....	11
4. SITES D'ACTIVITES HISTORIQUES.....	11
4.1. DEFINITION.....	11
4.2. LOCALISATION	11
4.3. ATTRIBUTS.....	11
5. ZONES ECONOMIQUES URBAINES.....	12
5.1. DEFINITION.....	12
5.2. GEOMETRIE	12
5.3. ATTRIBUTS.....	14
6. POLES ECONOMIQUES	15
6.1. DEFINITION.....	15
6.2. DONNEES EXISTANTES	15
6.3. GEOMETRIE	15
6.4. CONFLITS AVEC LES DONNEES EXISTANTES	15
6.5. STRUCTURE DE LA TABLE	16
6.6. RENSEIGNEMENT DES ATTRIBUTS.....	17
6.6.1. Commune et code INSEE	17
6.6.2. Identifiant du pôle	17
6.6.3. Surface.....	17
6.6.4. Nombre de sites.....	17
6.6.5. Date et opérateur de saisie	17
6.6.6. Référence spatiale et précision géométrique	18

1. Identification du document

1.1. Informations générales

Projet	Constitution de bases de données géographiques sur l'activité économique et l'occupation du sol en Picardie.
Titre du document	Constitution d'une base de données économique 2010. Méthodologie.
Version	2
Date	30/07/2013
Auteurs	SIRS
Destinataires	GéoPicardie

1.2. Historique des évolutions

Version	Date	Objet de la version	Auteur
1	19/06/2013	Méthodologie zone test	Sylvain DEBORDE
2	30/07/2013	Méthodologie globale	Sylvain DEBORDE

1.3. Circuit de validation

Version	Validation contractuelle
1	Nom : Lionel MEQUIGNON Organisme : SIRS Fonction : responsable projets occupation du sol Date : 19/06/2013
2	Nom : Cédric ANSARD Organisme : Conseil Régional de Picardie Direction de la Prospective, de la Programmation, de l'Aménagement Durable et de l'Europe Fonction : Responsable du Département SIG, ressources et connaissance territoriale Date : 26/06/2013

2. Introduction

Le présent document détaille la méthodologie de création d'une base de données économique localisant les zones d'activité sur l'ensemble de la région Picardie.

Ce travail a été réalisé en corrélation avec la mise à jour 2010 de l'occupation du sol régionale 1992-2002.

Bien que les couches produites ne distinguent au final que les sites (couche SITE) et les pôles (couche POLE) la méthode est présentée selon le détail des types de zones envisagées :

- Les zones d'activités industrielles et commerciales,
- Les sites d'activités historiques,
- Les zones économiques urbaines,
- Les pôles économiques.

3. Zones d'activités industrielles et commerciales

3.1. Définition

Ce sont de grands ensembles de commerces, d'industries ou d'activités. Ils sont situés principalement en périphérie des agglomérations (image ci-dessous).



ZI du Meret et ZA de la ferme des sables, à Breuil-le-Sec.

3.2. Géométrie

3.2.1. Données existantes

Certains acteurs locaux ont déjà mis en œuvre une numérisation de ces zones :

PNR Oise-Pays de France	donnée « sites_pnr_olv_aduga_arc »
Agence d'urbanisme et de développement de la vallée de l'Oise (Oise la Vallée)	
Agglomération Du Grand Amiénois (ADUGA)	
Agglomération de la Région de Compiègne (ARC)	
CCI Oise	
CCI Littoral Normand Picard	

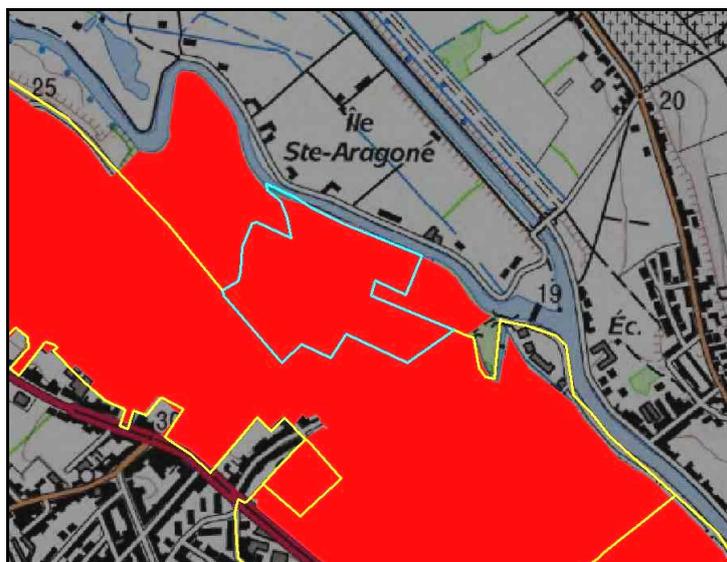
La donnée « sites_pnr_olv_aduga_arc » a été reprise telle quelle.
Elle a également servi de référence pour la structure de la base de données finale.

Les données fournies par les CCI de l'Oise et du Littoral Normand Picard ne sont pas totalement renseignées et présentent quelques soucis de topologie (superpositions et trous).
Elles ont subi un nettoyage de leur géométrie (outil « Intégrer » d'ArcGis 9.3) avec une tolérance de 5 m.

Leur géométrie a été utilisée lorsqu'il n'y a pas d'intersection avec les polygones de la donnée « sites_pnr_olv_aduga_arc » de référence.

Dans le détail :

- Ajout des polygones n'intersectant pas les polygones de référence,
- Suppression des polygones « compris dans » une zone tampon de 50 m autour des polygones de référence,
- Sélection des polygones restants, en intersection avec les polygones exogènes,
- Suppression des polygones compris ou quasiment compris dans les polygones de référence,
- Mise à jour des polygones à partir des polygones en intersection conservés.



Ci-dessus, un polygone (en surbrillance), intégré dans une zone de la donnée « sites_pnr_olv_aduga_arc » (en rouge). Fond scan 25 IGN.
→ il n'est pas conservé.



Ci-dessus, un polygone (en surbrillance) intersectant un polygone issu de la donnée de référence (en bleu). Fond scan 25 IGN.
→ ces contours sont utilisés pour créer un nouveau polygone dans la zone extérieure.

3.2.2. Données sources supplémentaires

Les données Navteq, dans l'emprise de l'Occsol 2010 et les polygones de cette occsol sont utilisés pour compléter le repérage de ces zones.

3.2.2.1. Occsol 2010

En ce qui concerne l'Occsol 2010, à l'exception des 1212 : « traitement et production d'eau » et des 1216 : « friches industrielles ou urbaines » les polygones codés en 121 : « zones industrielles et commerciales » sont utilisés.

Les codes 1211 : « emprises administratives » sont repris pour isoler les établissements hospitaliers. Ils ont été repérés par intersection avec les thèmes santé de la BD Topo et des hôpitaux repérés dans les données Navteq.

Un tri a ensuite été réalisé des polygones isolés de plus de 2 ha ou des polygones contigus dont la surface cumulée est supérieure à 2 ha.

Dans le détail :

- Fusion des polygones de 121 de l'Occsol 2010,
- Sélection des polygones résultants de plus de 2 ha,
- Sélection des polygones de l'Occsol 2010 initial compris dans ces polygones de plus de 2 ha.

Ensuite, de la même manière que les données CCI fournies ont été ajoutées à la donnée de référence « sites_pnr_olv_aduga_arc », la géométrie de ces polygones a été utilisée à défaut de données existantes correspondantes.

Une différence cependant : la sélection des polygones « compris dans » a été réalisée à partir de l'emprise des données existantes étendues de 50 m.

3.2.2.2. Données Navteq

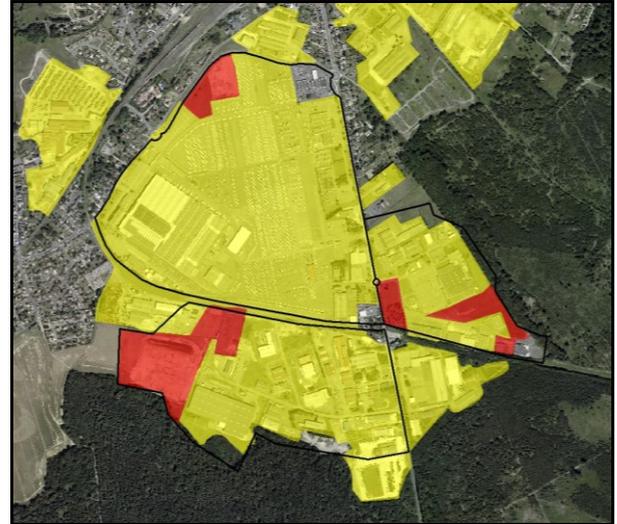
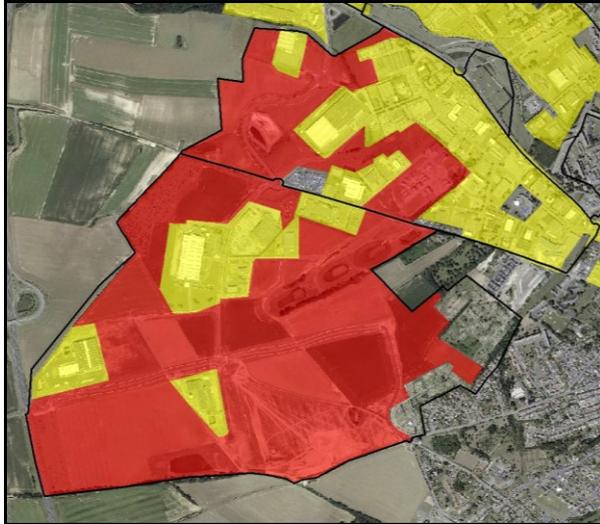
Les données Navteq sont géométriquement peu précises : des décalages de plusieurs dizaines de mètres sont constatés localement avec l'orthophotographie.

Elles ont alors été utilisées à défaut de données existantes ou de polygones issus de l'Occsol 2010 correspondants.

3.2.3. Cas des polygones délimités par plusieurs types de données

3.2.3.1. Données Navteq et Occsol

Ces cas sont fréquents dans les zones sans données existantes autres.



Zones issues de l'occsol (en jaune) et de Navteq (en rouge),
avec les contours de la donnée Navteq brute en noir).

Dans ces situations, les polygones issus de l'occsol et de Navteq sont fusionnés puis découpés suivant les contours de la donnée Navteq initiale (trait noir).

En pratique, le travail se fait en affectant manuellement l' « id_navteq » correspondant aux polygones issus de l'occsol puis par une fusion selon cet « id ». Des découpes sont réalisées le cas échéant.



Délimitation finale des zones (en traits bleus) après fusion
suivant les contours Navteq initiaux.

La source est précisée au final dans le champ SOURCE = « occsol, navteq dissolve »

3.2.3.2. Navteq, Occsol et données existantes



Limite des zones Navteq initiales (en trait noir), polygones issus de l'occsol (en jaune), des données navteq (en rouge) et des données CCI (en bleu).

On constate dans l'image ci-dessus que le découpage Navteq ne coïncide pas avec le découpage des données CCI.

Dans ce cas, la priorité est donnée au découpage des données existantes (ici les données CCI), elles ne seront en aucun cas fusionnées ou modifiées.

Dans un second temps les zones issues de la fusion des données occsol et Navteq sont découpées suivant les limites Navteq initiales (trait noir).

3.3. Structure de la table

	Nom du champ	Descriptif
Tronc commun	IDSITE	Identifiant du site
	IDPOLE	Identifiant du pôle
	SITE_NOM	Nom du site
	SITE_ETAT	Etat du site (existant, extension création)
	SURF_BRT	Surface totale du site (en m ²)
	SURF_NET	Surface cessible du site (en m ²)
	DATE_SAI	Date de saisie
	DATE_MAJ	Date de mise à jour
	OP_SAI	Personne ayant saisi la mise à jour
	ORG_SAI	Organisme dont dépend l'opérateur de saisie
	REF_SPA	Source d'information spatiale
	PREC_QTV	Dénominateur de l'échelle
Champs ajoutés	TYPE_ZONE	Type de zone (zone d'activités, zone économique urbaine, site isolé)
	SOURCE	Donnée source de la géométrie du polygone
	NBRE_ETS	Nombre de points établissements compris dans le polygone
	NB_TRANS	Nombre d'établissements de logistique et transport compris dans le polygone
	NB_TERTSUP	Nombre d'établissements tertiaires supérieurs compris dans le polygone
	NB_MED	Nombre d'établissements médicaux et de santé compris dans le polygone
	NB_INDUS	Nombre d'établissements industriels compris dans le polygone
	NB_COMM	Nombre d'établissements commerciaux compris dans le polygone
	NB_BTP	Nombre d'établissements de construction et BTP compris dans le polygone
	NB_TERTAU	Nombre d'établissements tertiaires autres compris dans le polygone
	NB_ADMIN	Nombre d'établissements administratifs compris dans le polygone
	EFF_TRANS	Somme des effectifs des établissements de transport et logistique compris dans le polygone
	EFF_TERSUP	Somme des effectifs des établissements tertiaires supérieurs compris dans le polygone
	EFF_MED	Somme des effectifs des établissements médicaux et de santé compris dans le polygone
	EFF_INDUS	Somme des effectifs des établissements industriels compris dans le polygone
	EFF_COMM	Somme des effectifs des établissements commerciaux compris dans le polygone
	EFF_BTP	Somme des effectifs des établissements de BTP compris dans le polygone
	EFF_TERTAU	Somme des effectifs des établissements tertiaires autres compris dans le polygone
	EFF_ADMIN	Somme des effectifs des établissements administratifs compris dans le polygone
EFF_TOTAL	Effectif total des établissements compris dans le polygone	

3.4. Renseignement des attributs

3.4.1. Identifiants du site

L'IDSITE est construit de la manière suivante :
CODE INSEE de la commune + codage à deux lettres

Dans le détail :

- Un identifiant de zone (IDZONE) est calculé, afin de pouvoir repérer les zones individuellement dans les étapes suivantes,
- La couche des communes, issue de la BD Topo, et la couche des zones d'activité on fait l'objet d'une Union,
- La table de la couche résultante a été triée selon l>IDZONE et selon la surface décroissante des polygones résultants,

- Un récapitulatif du code INSEE correspondant à la surface la plus élevée (commune dans laquelle un site donné a la plus grande surface) a ensuite été réalisé sur le champ IDZONE, De cette manière, dans le cas d'un polygone partagé entre plusieurs communes, le code INSEE retenu est celui de la commune où le site a la plus grande surface.
- La table des sites a ensuite été triée par code INSEE et par surface décroissante,
- Les sites déjà codés sont exclus de la table de calcul,
- Un traitement de la table exportée sous Excel a permis, par code INSEE, d'attribuer les deux lettres supplémentaires dans l'ordre de surface décroissante. De manière à distinguer aisément les identifiants calculés de ceux déjà affectés aux données existantes, le codage en lettre des nouveaux identifiants débute à « CA »,
- Le résultat de cette manipulation dans le tableur est joint, par l'intermédiaire de l'IDZONE, à la couche des sites initiaux pour renseigner l'IDSITE.

3.4.2. Identifiant du pôle

L'IDPOLE est construit de la manière suivante :
CODE INSEE de la commune + codage sur trois chiffres

Le détail du calcul de cet identifiant est présenté dans le chapitre 6.

3.4.3. Nom du site

Le nom du site est extrait des données existantes utilisées.

3.4.4. Surface

La surface est calculée en m² pour les polygones non déjà renseignés dans les données existantes.

3.4.5. Saisie

La date, l'opérateur et l'organisme de saisie sont renseignés comme suit :
19/06/2013, Sylvain Deborde, SIRS

3.4.6. Référence spatiale et précision géométrique

La référence spatiale est extraite des données existantes.
Pour les polygones non renseignés de cette manière, les sources « navteq » et « occsol » constituent également des références spatiales.

La précision géométrique est renseignée par l'échelle de référence.
Pour l'Occsol 2010, l'échelle de production de la donnée a été utilisée.
Pour les autres sources, non renseignées dans les données existantes, elle a été estimée en regard des décalages constatés avec les orthophotographies.

3.4.7. Nombre d'entreprises

Le nombre d'entreprise par site a été calculé par l'intermédiaire d'une jointure spatiale entre la couche des établissements et les polygones de sites.

Plusieurs récapitulatif successifs par IDSITE, après des sélections par types d'activité, permettent d'obtenir par site le nombre total d'établissements et celui de chaque type d'activité considéré.

3.4.8. Effectifs

Les effectifs par site de chaque secteur d'activité ont été calculés par l'intermédiaire d'une jointure spatiale entre la couche des établissements et les polygones de sites.

Plusieurs récapitulatif successifs par IDSITE, après des sélections par types d'activité, permettent d'obtenir les effectifs cumulés de chaque type d'activité considéré.

4. Sites d'activités historiques

4.1. Définition

Sites souvent anciens, non intégrés dans le tissu urbain, sans réelles politiques d'aménagement.
Exemple : site Continental à Clairoux (image ci-dessous).



ZI Le bac à l'aumone : site Continental, à Clairoux.

4.2. Localisation

En dehors des données « sites_pnr_olv_aduga_arc » dans lesquels un type « établissement isolé » est renseigné, les sites historiques ont été repérés à l'aide des données Navteq.

Les sites n'intersectant pas les zones industrielles, commerciales, les aéroports ou les emprises d'établissements de santé de la couche Navteq ont été considérés comme isolés.

Il s'agit évidemment d'une approximation dépendante des limites de la donnée Navteq : précision géométrique, actualisation, exhaustivité.

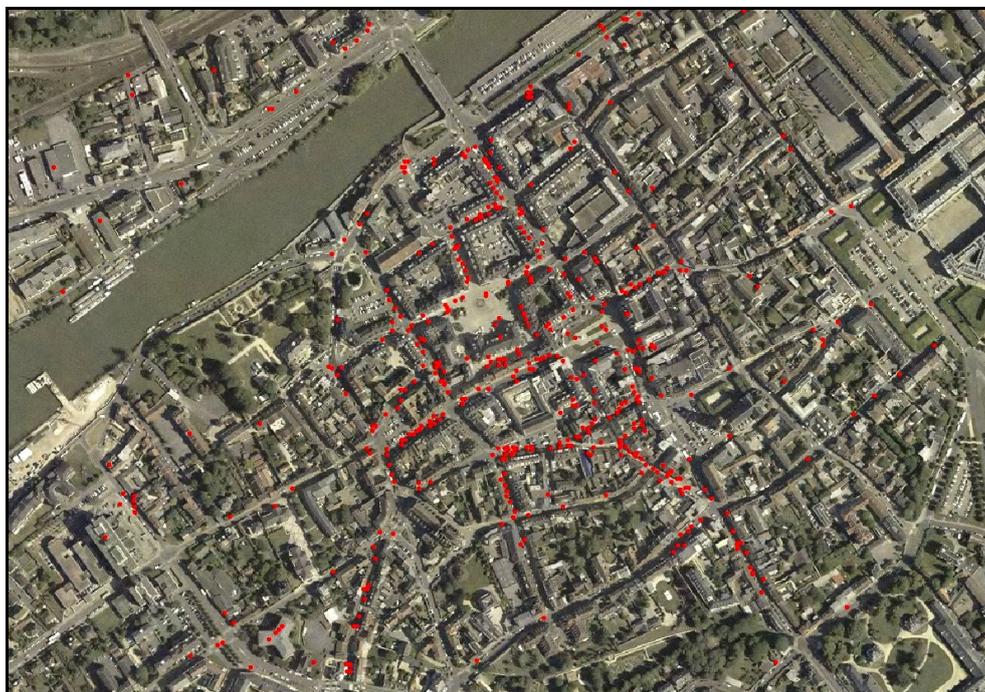
4.3. Attributs

Ces sites isolés sont repérés dans la couche SITES par le champ « TYPE_ZONE » renseigné comme « sites isolés ».

5. Zones économiques urbaines

5.1. Définition

Ces zones économiques sont liées aux services et commerces disponibles dans les agglomérations. Un certain nombre d'activités peuvent être considérées comme typiques des « centre ville » et sont repérées dans la couche ponctuelle de localisation des établissements.



Centre ville de Compiègne, avec concentration d'établissements typiques (en rouge).

5.2. Géométrie

La géométrie de ces zones a été créée par érosion-dilatation autour :

- des points correspondant aux établissements de type « centre ville »,
- des mairies, localisées à l'aide de la BD Topo IGN.

Une zone tampon de 50 m, avec fusion, a été générée autour des points considérés.

Pour des raisons de capacité de calcul, cette zone tampon a été générée en format raster, avec une précision de 10 m.

Les polygones obtenus ont ensuite été rognés de 20 m.

On obtient ainsi une zone d'influence d'un rayon de 30 m, agglomérant les zones d'influence de chaque point.

Les polygones résultants ont été lissés (tolérance de 100 m utilisée dans l'outil « lisser un polygone » d'ArcGis 9.3).



Résultat de l'érosion-dilatation (en rouge) à partir de la sélection des points établissements de type « centre ville » (en jaune). Fond scan 25 IGN.

Dans un deuxième temps les polygones n'étant pas en intersection avec l'urbain de l'Ocssol 2010 (code 11) et ceux qui ont une surface de moins de 1ha ont été éliminés.



Résultat du tri des polygones supérieurs à 1 ha.

En dernière étape, le croisement entre ces polygones et les polygones de zones d'activités a été réalisé. La couche des zones économiques urbaines a subi un « Erase » par la couche des zones d'activités (image ci-dessous).

Les discontinuités apparaissant alors localement (image ci-dessous) ont été comblées manuellement.

Reque : plusieurs soucis de continuité des polygones de la couche « sites_pnr_olv_aduga_arc » ont été repérés à cette occasion.

Les polygones résultants inférieurs à 1 ha ont été traités manuellement : la plupart ont été éliminés mais certains, jugés relativement structurants, ont été conservés.



Superposition de la couche des zones d'activités (en jaune)
et des zones économiques urbaines (en rouge).
Un trou (flèche verte) apparaît ici entre les deux couches.

5.3. Attributs

Ces zones économiques urbaines sont repérées dans la couche SITES par le champ « TYPE_ZONE » renseigné comme « zones économiques urbaines ».

Pour ces zones, le nombre d'établissements total (NBRE_ETS) et de type administratif (NB_ADMIN) a été obtenu à l'aide d'une jointure spatiale entre la couche des points établissements de type « centre ville » et les polygones de zones économiques urbaines.

Deux récapitulatifs sur l'identifiant des polygones ont été réalisés :

- un récapitulatif global,
- un second après sélection des établissements de type administratif.

Une jointure avec la couche initiale des polygones a ensuite permis de renseigner ces deux champs.

6. Pôles économiques

6.1. Définition

Les pôles regroupent plusieurs sites d'activités contigus ou géographiquement proches.

6.2. Données existantes

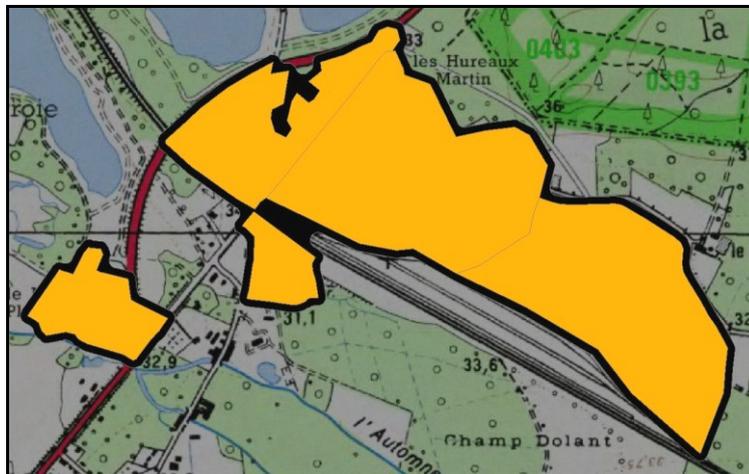
La couche « sites_pnr_olv_aduga_arc » possède déjà certains « idpole » renseignés. Ils permettent de constituer quelques pôles en amont de la création par la seule géométrie.

6.3. Géométrie

Les limites de ces pôles ont été obtenues par création d'une zone tampon de 20 m, avec fusion, autour des polygones de la couche des zones d'activités.

Cette zone tampon a subi un lissage par génération d'une zone tampon à 20 m puis – 20 m, sans fusion.

Les zones économiques urbaines n'y ont pas été intégrées.



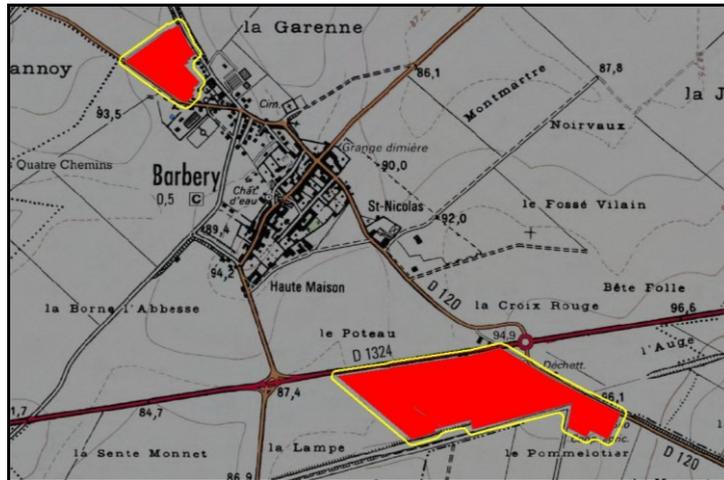
Exemples de pôles (en noir) générés à partir des polygones de zones d'activités (en jaune)

6.4. Conflits avec les données existantes

Dans quelques cas, les polygones obtenus par cette zone tampon de 20 m englobent un ou plusieurs pôles résultant des « idpole » renseignés dans les données existantes.

De la même manière, plusieurs polygones distincts issus de la génération de cette zone tampon peuvent appartenir à un pôle unique dans la donnée existante.

La géométrie a alors été mise à jour pour correspondre aux attributs initiaux.



Les deux polygones rouges ci-dessus appartiennent à un pôle initial unique dans les données existantes.
Les deux pôles générés par zones tampons (en jaune) ont donc été combinés.



Les polygones roses et verts appartiennent à des pôles distincts dans la donnée existante.
Le pôle généré par zones tampons (en jaune) a donc été découpé (flèche rouge).

6.5. Structure de la table

Nom du champ	Descriptif
IDPOLE_INI	Identifiant initial du site (donnée « sites_pnr_olv_aduga_arc »)
IDPOLE_NEW	Identifiant calculé
NOM_POLE	Nom du pôle
INSEE	Code INSEE de la commune
COMMUNE	Nom de la commune
SURF_BRT	Surface totale du pôle
SURF_NET	Surface cessible du pôle
SURF_DISP	Surface disponible à la vente sur le pôle
NB_SITES	Nombre de sites inclus dans le pôle
DATE_SAI	Date de saisie
DATE_MAJ	Date de mise à jour
OP_SAI	Opérateur de saisie
ORG_SAI	Organisme dont dépend l'opérateur de saisie
REF_SPA	Source d'information spatiale
PREC_QTV	Dénominateur de l'échelle

6.6. Renseignement des attributs

6.6.1. Commune et code INSEE

Le code INSEE est celui de la commune ou se situe le pôle.

Dans le cas d'un polygone partagé entre plusieurs communes, le code INSEE retenu est celui de la commune ou le pôle a la plus grande surface.

Dans le détail :

- Un identifiant de zone (IDZONE) a été calculé, afin de pouvoir repérer les pôles individuellement dans les étapes suivantes,
- La couche des communes issues de la BD Topo et la couche des pôles ont fait l'objet d'une « Union »,
- La table de la couche résultante a été triée selon l'IDZONE et selon la surface décroissante des polygones résultants,
- Un récapitulatif du code INSEE correspondant à la surface la plus élevée (commune dans laquelle un site donné a la plus grande surface) a ensuite été réalisé sur le champ IDZONE.

Rque : quelques incohérences ont été constatées entre l'IDPOLE déjà renseigné pour certains polygones des données existantes et le code INSEE ainsi calculé.

6.6.2. Identifiant du pôle

Les IDPOLE renseignés dans les données existantes ne sont pas homogènes. La majorité sont construits de la manière suivante :

Code INSEE de la commune + codage à 3 chiffres

Un nouvel identifiant est calculé seulement pour les polygones dont l'IDPOLE n'est pas déjà renseigné dans les données existantes.

Quelques incohérences ont été constatées entre le code INSEE calculé et celui saisi dans la base initiale. De plus, les IDPOLE renseignés dans les données existantes ne sont pas codés de manière homogène.

Il a alors été choisi de séparer les identifiants en deux champs : IDPOLE_INI et IDPOLE_NEW.

Pour calculer l'IDPOLE_NEW, les étapes suivantes ont été nécessaires :

- La table des pôles a été triée par code INSEE et par surface décroissante,
- Un traitement de la table exportée sous Excel a permis, par code INSEE, d'attribuer les trois chiffres supplémentaires dans l'ordre de surface décroissante,
- Le résultat de cette manipulation dans le tableur est joint, par l'intermédiaire de l'IDZONE, à la couche des pôles initiaux pour renseigner l'IDPOLE_NEW.

6.6.3. Surface

La surface brute en m² est calculée par le SIG.

6.6.4. Nombre de sites

Le nombre de sites compris dans chaque pôle est obtenu par l'intermédiaire d'une jointure spatiale entre la couche des pôles et celle des sites.

Un récapitulatif par IDZONE permet d'obtenir le nombre de sites.

6.6.5. Date et opérateur de saisie

La date, l'opérateur et l'organisme de saisie sont renseignés comme suit :

04/07/2013, Sylvain Deborde, SIRS

6.6.6. Référence spatiale et précision géométrique

La multitude des sources utilisées pour la création des polygones de sites d'activité rend délicate l'appréciation de la précision géométrique du contour des pôles.

Il a donc été choisi d'utiliser l'échelle de référence de 1/10000^{ème} : celle des données Navteq, les moins précises géométriquement.