

IDENTIFICATION DES SERVICES ET CALCUL D'UN INDICE DE POLARITE



Liberté • Égalité • Fraternité

RÉPUBLIQUE FRANÇAISE

**Direction Départementale
des Territoires**

COMMANDITAIRE : DDT 31

ENCADRANTS : SEBASTIEN LE CORRE, MARTIN PAEGELOW, DDT 31

Table des matières

Introduction	5
1) Présentation du sujet	6
1.1) Mots-clés liés à la polarité.....	6
1.2) Objectifs opérationnels.....	6
1.3) Contraintes de définition de l'indice de polarité	8
2) Inventaire des données	9
3) Données non utilisées	15
3.1) Echelle - Données carroyées	15
3.2) Emploi.....	18
3.2.1) Nombre d'établissements employeurs du secteur privé et effectifs salariés par commune - Base SEQUOIA (ACOSS)	18
3.2.2) Activité des résidents – Base IRIS (INSEE)	18
3.2.3) Les stocks d'établissements et effectifs salariés par commune –Répertoire des Entreprises et des Etablissements (INSEE)	19
3.3) Démographie.....	20
3.4) Equipements	20
3.5) Mobilité et transport.....	21
3.5.1) Mobilités scolaires et professionnelles des individus : déplacements commune de résidence / commune de scolarisation – Fichiers détails (INSEE).....	21
3.5.2) Flux > 100 des mobilités scolaires et professionnelles - Base flux de mobilité (INSEE).....	22
3.5.3) Thèmes réseaux routier et ferroviaire - BD TOPO (IGN)	22
3.6) Caractérisation des surfaces qui ne sont pas dans les carreaux INSEE.....	23
4) Données utilisées	25
4.1) Indicateur activité - Implantation d'établissements - Base SIRENE.....	25
4.1.1) Présentation	25
4.1.2) Application	28
4.1.3) Critique.....	28

4.2) Indicateur démographique - Estimation de la population	28
4.2.1) Présentation	28
4.2.2) Application	31
4.2.3) Critique	32
4.3) Indicateur d'équipement - Implantation d'équipements	33
4.3.1) Présentation	33
4.3.2) Application	36
4.3.3) Critique	37
5) Paramètres du plugin « Carte de chaleur » de QGIS	37
5.1) Le Rayon	38
5.2) Taille de la cellule	39
5.3) Forme du noyau	42
5.4) Pondération :	44
6) Elaboration de l'indice de polarité.....	44
6.1) Disposer de données sources ponctuelles	44
6.2) Ajout de 4 points à chaque couche	44
6.3) Traitement de chaque nuage de points par le plugin "carte de chaleur"	44
6.4) Alignement des rasters	44
6.5) Suppression des valeurs -9999.....	44
6.6) Addition des rasters (calculatrice raster.)	45
6.7) Symbologie	45
7) Retour critique sur le projet.....	50
8) Résultats et limites	51
9) Perspectives	52
10) Conclusion.....	52
11) Annexes	53

Table des illustrations

Figure 1 : Critères de définition de l'indice de polarité	8
Figure 2 : Inventaire des données page 1	10
Figure 3 : Inventaire des données page 2	11
Figure 4 : Arbre décisionnel - Indice de polarité	14
Figure 5 : Carte traitement des mailles	16
Figure 6 : Carte des données en dehors des mailles INSEE	24
Figure 7 : Carte de présentation des données SIRENE	27
Figure 8 : Carte de présentation des données démographiques	30
Figure 9 : Carte de présentation des données d'équipement	35
Figure 10 : Carte de test du rayon	39
Figure 11 : Carte de test de la résolution	41
Figure 12 Illustration des fonctions utilisées pour calculer une densité de noyau	42
Figure 13 : Carte de test du type de noyau	43
Figure 14 : Onglet style de la carte des polarités de la Haute-Garonne	45
Figure 15 Onglet style de la carte des polarités du SCoT Grande agglomération Toulousaine	46
Figure 16 : Carte finale Haute-Garonne	47
Figure 17 : Carte finale SCoT Grande Agglomération Toulousaine	48
Figure 18 : Paramètres cartes finales : réglages figurant dans le projet QGIS	49
Figure 19 : Tableau des apports et des limites du projet	50
Figure 20 : Carte de zonage en aires urbaines	53
Figure 21 : Carte de l'évolution de la population communale 1975 – 2015	54
Figure 22 : Carte du nombre d'établissement du secteur privé par commune	55
Figure 23 : Carte d'évolution de la population entre 2010 et 2015 dans le carroyage INSEE	56

Introduction

Aujourd'hui la Haute-Garonne constitue l'un des départements français les plus dynamiques sur le plan démographique avec environ 91 000 nouveaux habitants entre 2010 et 2015. Toulouse Métropole a absorbé à elle seule près de la moitié de cet afflux et totalise en 2015 environ 755 000 habitants, soit 56 % de la population départementale.

La métropole affirme ainsi de plus en plus son influence sur les territoires avoisinants, ce qui ne se traduit pas uniquement dans une dimension démographique. On assiste à un phénomène de concentration spatiale des fonctions urbaines et de diversification au sein d'un nouveau système productif.

Cette tendance s'inscrit dans le cadre d'un processus global de métropolisation des territoires qui se traduit notamment par l'étalement urbain, l'accroissement des mobilités et de leur coût, la dévitalisation des centres-villes et la recomposition de l'ancrage local, et qui mobilise des nouveaux modes de gouvernance urbaine.

Acteur de l'aménagement du territoire et du développement durable, la Direction Départementale de la Haute-Garonne (DDT 31) cherche à analyser les dynamiques démographiques au regard des équipements existants (entreprises, commerces, santé, loisirs, etc...) et des activités.

Ses politiques et ses choix d'investissement constituent ainsi des leviers d'action pour orienter l'installation des populations dans les territoires disposant des meilleures conditions de services ou de déterminer l'implantation d'un équipement en tenant compte de l'organisation du territoire.

L'objectif de cette étude vise à concevoir une méthode d'identification des polarités présentes (démographie, équipements et activité) à l'échelle du département de la Haute-Garonne avec un mode de représentation au niveau infra-communal. Il s'agit d'une démarche exploratoire visant à identifier les données disponibles, la méthodologie employée, les biais éventuels, l'interprétation des résultats et les perspectives.

Son application pratique permettra de profiler une hiérarchie urbaine au sein du département afin de constituer une aide à la décision pour les aménageurs : selon l'intensité du secteur étudié dans la polarité, on peut par exemple évaluer s'il est apte à accueillir un équipement avec l'allocation de ressources publiques. L'indice développé contribuera ainsi à un aménagement du territoire équilibré en déterminant s'il est pertinent de maintenir, développer ou d'infléchir une polarité.

1) Présentation du sujet

1.1) Mots-clés liés à la polarité

L'indice de polarité mobilise différentes notions qui permettent de mieux cerner les enjeux et les objectifs de l'étude :

- **Nœud** : Lieu d'interface entre des axes appartenant à un ou plusieurs réseaux. Le degré de connexion et la position d'un nœud par rapport aux autres détermine sa capacité de développement et son niveau de hiérarchisation, qui peut varier du niveau local au niveau mondial.
- **Pôle** : Lieu d'attraction des flux de personnes et de biens par recours au service le plus proche. Le poids et le degré de rareté des services détermine la hiérarchie des pôles et de la dimension de leur aire d'influence ou aire de chalandise.
Le rayonnement d'un pôle répond moins à une logique de distance (topographique) avec la proximité qu'à une logique de réseau (topologique) avec l'accessibilité. Ainsi les pôles peuvent entretenir davantage de liens entre eux qu'avec leur périphérie.
- **Centres** : Ils constituent le cœur de l'espace urbain où se concentrent les fonctions urbaines caractéristiques. Ils sont également hiérarchisés selon le volume de l'offre en service et par leur diversité. De façon générale, l'espacement des centres augmente et leur nombre diminue à mesure qu'ils disposent d'un niveau de service important.
- **Armature urbaine** : Structure de la hiérarchisation urbaine d'un territoire correspondant à l'agrégation des nœuds, des pôles et des centres. L'indice de polarité permet de retranscrire l'armature urbaine selon des critères de classification utilisés (démographie, activité, équipement, etc).
- **Bassin de vie** : Selon l'INSEE, « le bassin de vie constitue le plus petit territoire sur lequel les habitants ont accès aux équipements et services les plus courants. Le zonage en bassins de vie apporte un complément à travers l'analyse de la répartition des équipements et de leur accès ».
Il peut également être défini par la DATAR comme un « territoire présentant une cohérence géographique, sociale, culturelle et économique, exprimant des besoins homogènes en matière d'activités et de services ».

1.2) Objectifs opérationnels

La commande initiale du commanditaire décline des objectifs opérationnels permettant de cadrer le rendu :

- Développer un indice de polarité

L'indice de polarité est un indicateur d'analyse territoriale couramment utilisé qui ne dispose pas d'une méthodologie spécifique mais dont les caractéristiques peuvent varier selon son application.

Celles-ci recouvrent notamment :

- Les critères :

L'indice de polarité est un outil agrégé à partir d'indicateurs thématiques. Si le nombre de thématiques retenues n'est pas défini, on retrouve principalement comme constituantes la démographie, l'activité, les équipements et le transport.

Ces quatre thématiques ont été sélectionnées préalablement par le commanditaire pour constituer l'indice de polarité.

- Le mode de représentation :

- Le territoire :

En fonction de sa finalité, l'indice de polarité peut s'appliquer à un territoire de dimension variable, dont l'emprise varie généralement du niveau régional au niveau local.

Dans le cadre du projet, la DDT 31 a retenu son périmètre d'intervention à savoir la Haute-Garonne comme territoire pour effectuer l'étude.

- L'échelle :

L'aire d'étude retenue peut elle-même être représentée selon différents degrés de précision correspondant à la résolution spatiale choisie. On retrouve généralement l'échelle communale (découpage administratif) et infra-communale (découpage statistique ou spatial) avec l'IRIS, la maille ou le point. Le commanditaire a exprimé le besoin d'obtenir une représentation fine au niveau infra-communal, sans spécifier si elle doit correspondre à la maille ou au point.

- La temporalité :

L'indice de polarité peut s'appliquer selon deux types de temporalités : l'étude diachronique qui intègre l'évolution des polarités dans le temps c'est à-dire sur une période et l'étude synchronique qui donne une vision statique en se focalisant sur un moment donné.

Une étude diachronique est privilégiée pour l'application de l'indice de polarité bien que ce critère ne soit pas impératif.

Il s'agira de déterminer quelles caractéristiques s'avèrent réalisables et pertinentes pour effectuer cette étude.

- Rendre l'outil adaptable et reproductible sur d'autres territoires

La méthodologie retenue pour le projet sera restituée sous la forme d'un mémoire technique. Celui-ci devra notamment retranscrire et justifier les choix ayant conduit à générer l'indice de polarité dans sa forme finale. De cette façon, il sera possible de s'approprier la méthode adoptée et de l'ajuster si nécessaire sans avoir à effectuer les mêmes raisonnements.

Les données sources utilisées pour le calcul de l'indice de polarité devront également être détaillées pour pouvoir les modifier selon le besoin exprimé. Leur sélection se fera selon leur accessibilité et leur disponibilité à l'échelle nationale afin de pouvoir reproduire l'indice de polarité selon différentes temporalités ou modes de représentation.

- Identifier et interpréter les polarités présentes en Haute-Garonne

L'indice devra finalement permettre de faire ressortir les polarités sur l'aire d'étude qu'il s'agira d'interpréter au regard de l'organisation du territoire et des dynamiques en cours. Ces résultats seront à nuancer selon les éventuels biais méthodologiques explicités au préalable.

1.3) Contraintes de définition de l'indice de polarité

En complément des objectifs et des demandes exprimées par la DDT 31, l'indice de polarité est tributaire d'un certain nombre de contraintes liées à la disponibilité des données sources utilisées pour son agrégation.

Celles-ci sont précisées dans le schéma ci-dessous :

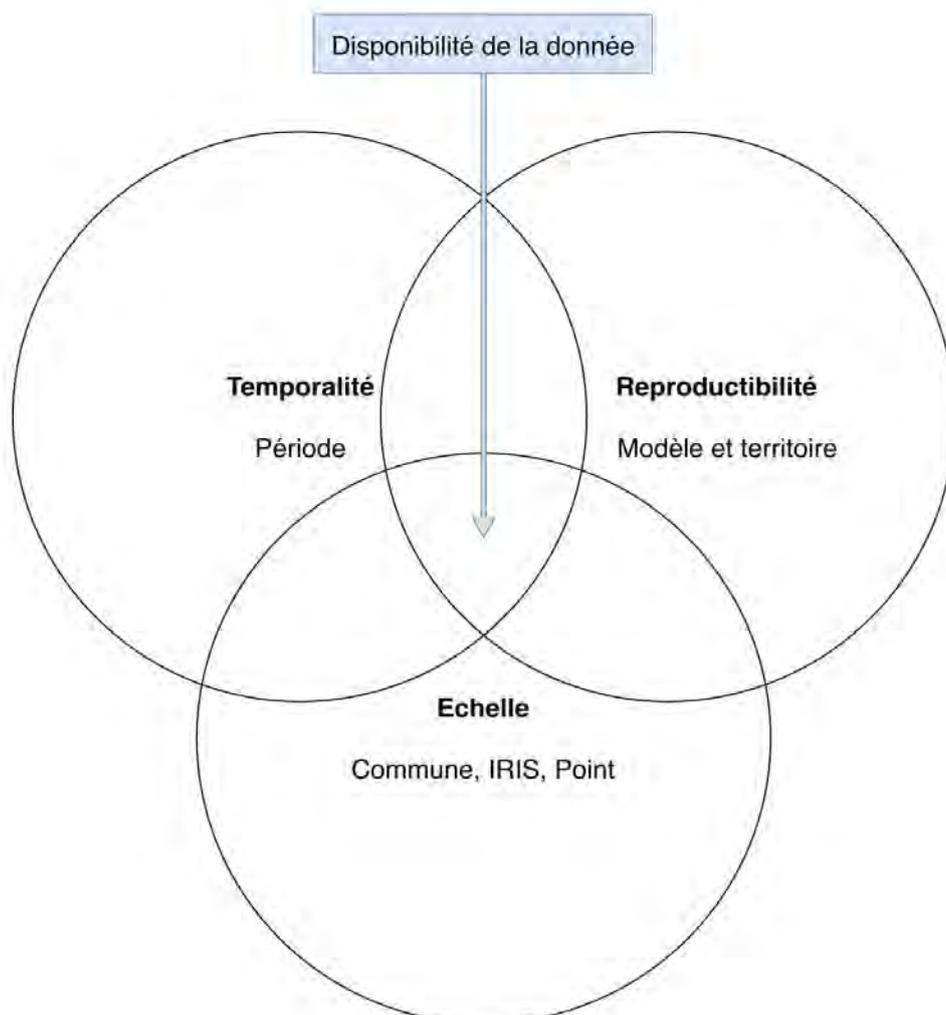


Figure 1 : Critères de définition de l'indice de polarité

Les trois critères que sont la temporalité, la reproductibilité et l'échelle dépendent de la donnée disponible à exploiter. Il s'agira de les satisfaire au maximum ces critères en tenant compte de cette variable.

2) Inventaire des données

Un inventaire des données disponibles a été effectué durant l'ensemble du projet pour accompagner le cheminement méthodologique. Il a été mis à jour suite aux différents apports de données des partenaires et à la redéfinition de certains objectifs.

La version finale du document correspond à un tableur au format Drive dont l'export figure ci-dessous :

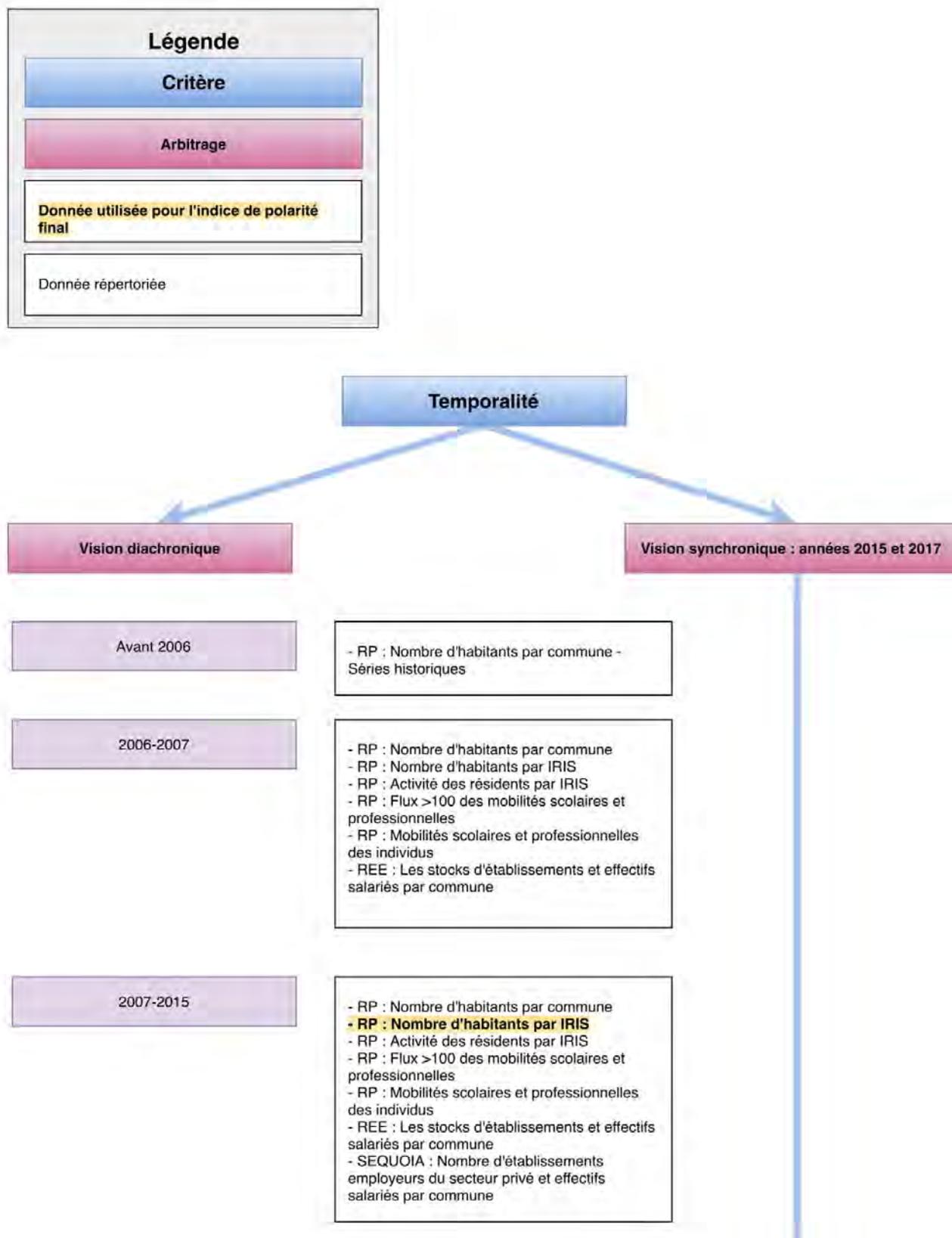
Données utilisées											
Thématique	Nom donnees	Nom détaillé	Fournisseur	Date	Source	Lien de téléchargement	Echelle	Territoire couvert	Géométrie	attribut géométrique	Observations
Emploi	Nombre et type d'établissements	Base Sirene des entreprises et de leurs établissements (SIREN, SIRET)	INSEE	2 décembre 2018 (MAJ quotidienne)	Base SIRENE	http://data.cquest.org/geo_sirene/last/	Toutes	France (avec filtrage Haute-Garonne)		csv avec x,y sur la plupart des équipements	Effectifs par classes d'intervalle
Emploi	Nombre et type d'établissements	Base Sirene des entreprises et de leurs établissements (SIREN, SIRET)	INSEE	6 mai 2018	Base SIRENE	https://data.haute-garonne.fr/explore/dataset/base-sirene/information/7disjunctive	Toutes	France			Effectifs par classes d'intervalle Pour les données SIRENE, autre source possible sur : https://www.sirene.fr/geo/publication
Emploi	Nombre et type d'établissements	Base Sirene des entreprises et de leurs établissements (SIREN, SIRET)	INSEE	2017-2019 (MAJ mensuelle)	Base SIRENE	https://www.data.gouv.fr/fr/datasets/base-sirene-des-entreprises-et-de-leurs-etablissements-siren-siret-fin-le-30-avril-2019/	Toutes	France			Effectifs par classes d'intervalle
Emploi	Nombre d'établissements employeurs du secteur privé et effectifs salariés par commune	Dénombrements au niveau NAF 732	ACOSS	2007 à 2017	Base SEQUOIA	https://www.acoss.fr/home/observatoire-economique/donnees-statistiques/bases-de-donnees.html	Commune	France (avec filtrage Haute-Garonne)			Pas de champ adresse
Emploi	Activité des résidents par IRIS	Recensement de la population - Base infracommunale (IRIS)	INSEE	2006-2015	RP	https://www.insee.fr/fr/statistiques/3627009#consulter	IRIS	France			Lien 2006-2014 : https://www.insee.fr/fr/recherche/recherche-statistiques?q=Activit%C3%A9%20des%20r
Emploi	Les stocks d'établissements et effectifs salariés par commune	Répertoire des entreprises et des établissements (REE) - Fichiers détail	INSEE	31 décembre 2006 à 2016	REE	https://www.insee.fr/fr/statistiques/2985296?sommaire=2538611#consulter	Commune	France (avec filtrage Haute-Garonne)			Effectifs par classes d'intervalle NB : on ne connaît pas la géographie utilisée pour le découpage communal
Démographie	Nombre d'habitants par IRIS	Recensement de la population - Base infracommunale (IRIS)	INSEE	2006-2015	RP	https://www.insee.fr/fr/statistiques?taille=100&debut=0&theme=1&categorie=3	IRIS	France			
Démographie	Nombre d'habitants par commune - Séries historiques	Recensement de la population - Base des principaux indicateurs	INSEE	1968, 1975, 1982, 1990, 1999, 2010 et 2015	RP	https://www.insee.fr/fr/statistiques/3565661#dictionnaire	Commune	France			
Démographie	Nombre d'habitants par commune	Recensement de la population - Base indicateurs de la fiche synthétique	INSEE	2006 - 2015	RP	https://www.insee.fr/fr/statistiques/3564100?sommaire=3561107	Commune	France	https://www.insee.fr/fr/statistiques/3564100?sommaire=3561107		Lien 2006-2014 : https://www.insee.fr/fr/information/2880845
Équipement	Dénombrement des équipements de commerce, services, santé, enseignement, sport-loisir et tourisme en 2017 (géolocalisé)	BPE (Base permanente des équipements)	INSEE	2017	BPE	https://www.insee.fr/fr/statistiques/3568638?sommaire=3568656#consulter-sommaire	Toutes	France entière	point	dbf avec x,y sur la plupart des équipements	Existe sur picto à l'échelle d'Occitanie !!!
Équipement	Etablissements scolaires 2nd degrés	Etablissements scolaires 2nd degrés de l'académie de Toulouse	Rectorat de l'académie (via picto)	extraction 2016	Base RAMESES	https://catalogue.picto-occitanie.fr/geosource/panierDownloadFronta?Parametage?LAYERIDTS=1704608	Toutes	Occitanie	point	shapefile	
Équipement	Etablissements scolaires du supérieur	Etablissements scolaires du supérieur de l'académie de Toulouse	Rectorat de l'académie (via picto)	2016	Base RAMESES	https://catalogue.picto-occitanie.fr/geosource/panierDownloadFronta?Parametage?LAYERIDTS=1705525	Toutes	Occitanie	point	shapefile	
Équipement	Etablissements scolaires	Etablissements scolaires de l'académie de Toulouse	Rectorat de l'académie (via picto)	2016	Base RAMESES	https://catalogue.picto-occitanie.fr/geosource/panierDownloadFronta?Parametage?LAYERIDTS=1006751	Toutes	Occitanie	point	shapefile	

Figure 2 : Inventaire des données page 1

Sur la base des données présentes dans l'inventaire, il a fallu effectuer un arbitrage sur les critères à retenir pour constituer l'indice de polarité, à savoir la temporalité, l'échelle et les thématiques des indicateurs. Pour y parvenir, nous avons procédé à une démarche "en entonnoir" où les critères ont été étudiés par ordre de complexité pour déterminer s'il était possible de les utiliser :

- Temporalité diachronique / synchronique : L'étude diachronique a été écartée car seule la donnée du nombre d'habitants par commune était accessible avant 2006 et des données indispensables telle que la BPE n'était accessible que pour l'année 2017 et la base SIRENE uniquement entre 2017 et 2019.
Un compromis a été trouvé en utilisant la donnée de population la plus récente, le nombre d'habitants par IRIS issu du recensement de la population de 2015 avec les données d'équipements et d'activité les plus anciennes, respectivement la BPE et la base SIRENE en version 2017.
NB : pour la base SIRENE, une version 2018 a finalement été retenue pour des raisons évoquées dans la suite du rapport.
- Echelle commune / point / IRIS : Pour ce critère, on peut voir que l'échelle comportant le plus de données correspond à la commune. Pour autant, la DDT 31 a souhaité privilégier une représentation fine au niveau infra-communal. L'échelle au point a été retenue car elle offre le plus de précision.
- Indicateurs thématiques démographie / activité / équipement : Les trois thématiques ont été retenus pour constituer l'indice de polarité. Cette décision a été prise en partenariat avec le commanditaire suite à la 1ere réunion de suivi intervenue le 23/01/19. La possibilité d'intégrer des données de transport et de mobilité à l'indice final a été écarté compte-tenu du temps de calculs importants et du délai imparti limité.

Ces arbitrages successifs ont été schématisés dans l'arbre décisionnel ci-dessous :



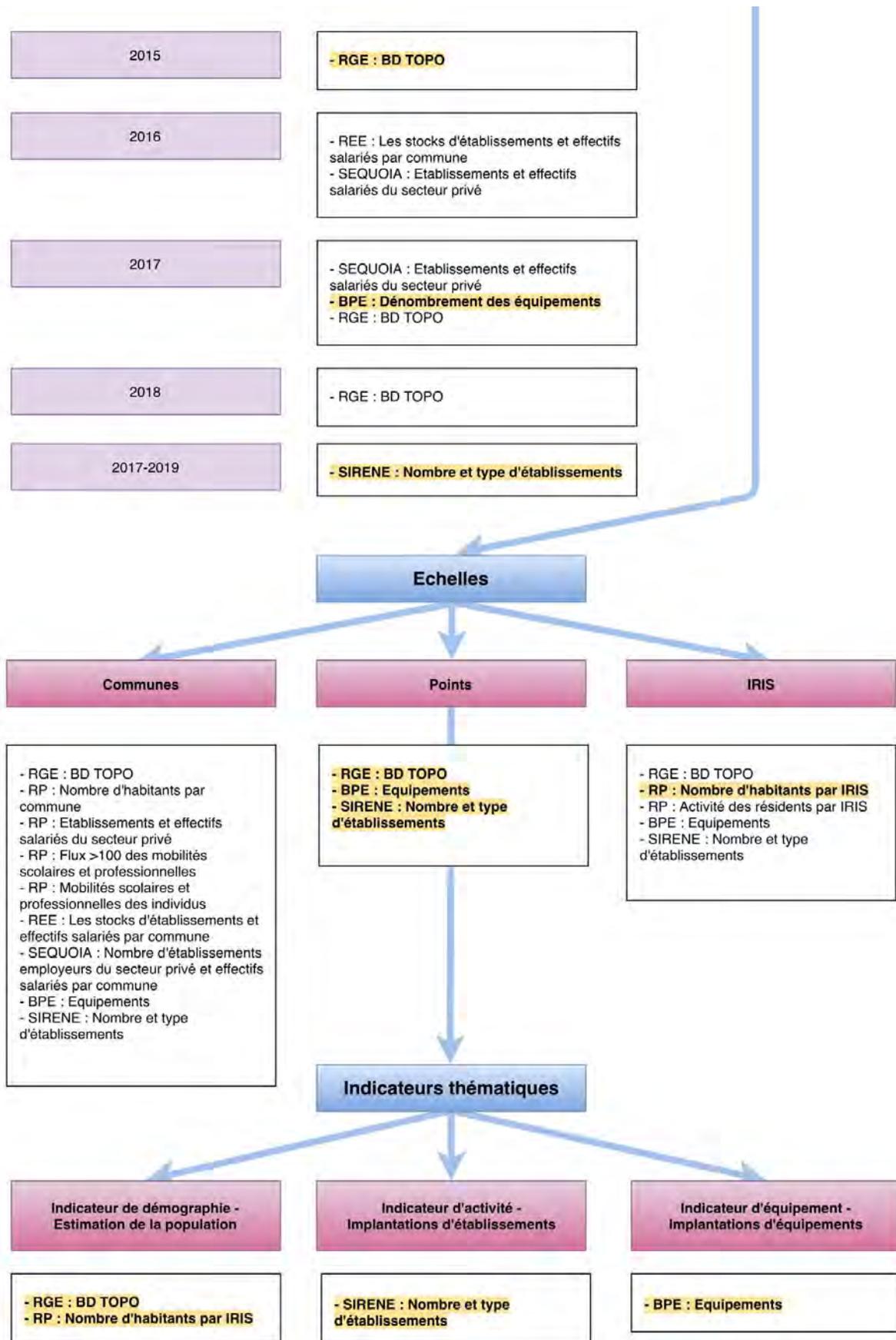


Figure 4 : Arbre décisionnel - Indice de polarité

3) Données non utilisées

La partie présente vise à justifier certains choix méthodologiques en présentant les principales données non utilisées et en effectuant un retour critique dans l'éventualité où elles seraient utilisées. Elle permet donc de comprendre pourquoi l'indice de polarité a été créé dans sa forme actuelle et quels ajustements sont envisageables au stade de l'intégration des données.

3.1) Echelle - Données carroyées

- Présentation

Les données carroyées proviennent de l'INSEE et correspondent à un découpage du territoire en mailles régulières et fines de 200 mètres de côté. Elles n'intègrent donc pas les limites administratives et les éléments topographiques.

Cette base a été réalisée en 2010 et comprend 18 variables sur la structure par âge des individus, les caractéristiques des ménages (locataire/propriétaire, etc.) et les revenus fiscaux au 31 décembre 2010.

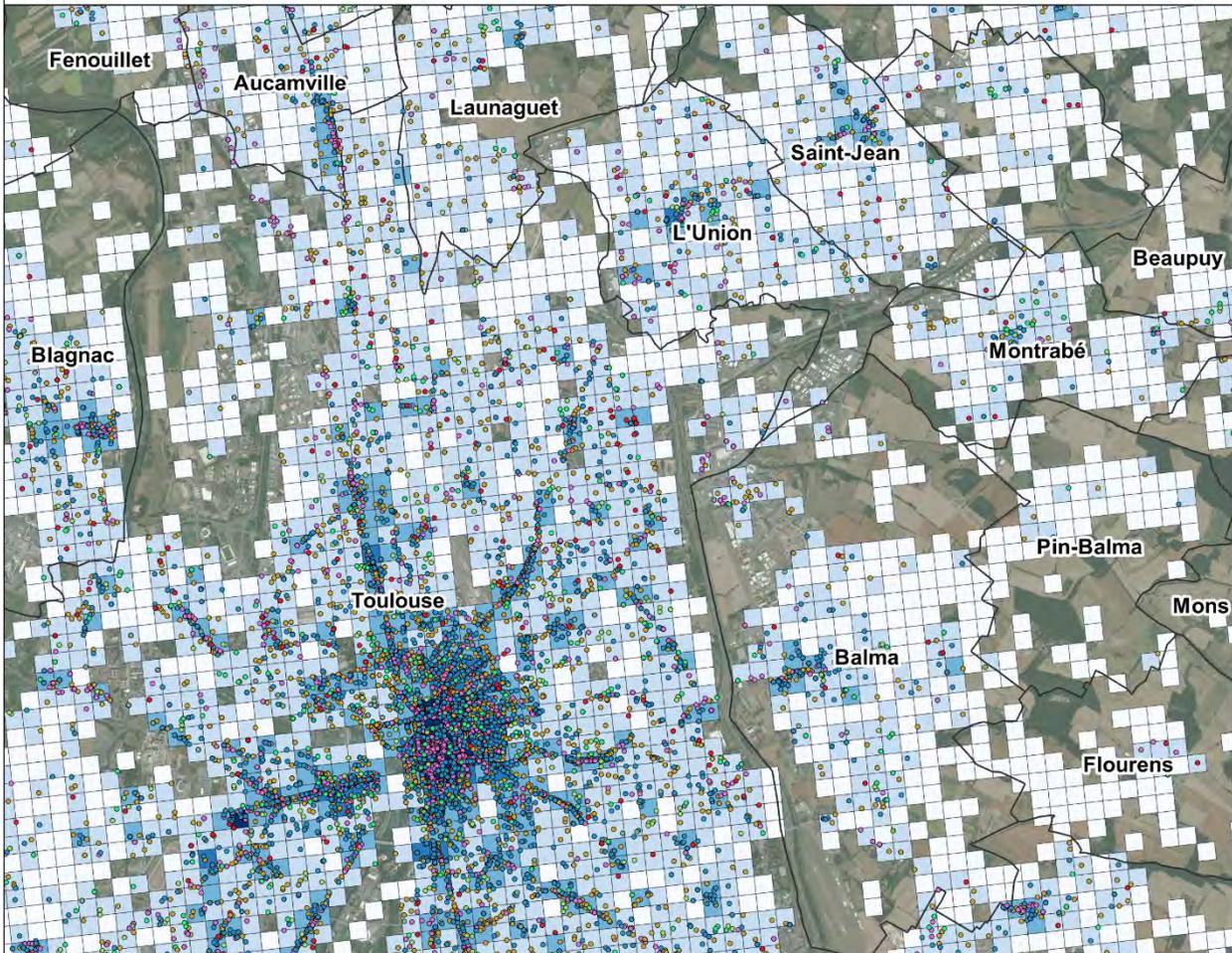
Ces données de source fiscale impliquent un traitement confidentiel : aucune information statistique - à l'exception du nombre total d'individus - n'est diffusée sur des carreaux de moins de 11 ménages. Cela ne constitue pas une limite car seule cette variable est nécessaire pour la réalisation de l'indice de polarité.

Ces carreaux de moins de 11 ménages sont regroupés avec des d'autres carreaux adjacents en rectangles de taille variable jusqu'à ce qu'ils soient supérieurs au seuil de 11 ménages.

A noter que les carreaux ne comprenant aucun ménage ne sont pas représentés, le maillage est donc incomplet : ainsi moins le secteur d'étude est peuplé, moins le carroyage sera couvrant en tenant compte d'une résolution de 200 mètres.

Ainsi même la région de Toulouse comprend un maillage partiel en raison de la présence de grandes infrastructures (aéroports de Blagnac, axes de communication), des zones d'activités, des espaces agricoles etc... On peut visualiser cette situation dans la carte ci-dessous :

Traitement des mailles
Illustration d'un indicateur : nombre total d'équipements
dans les environs de Toulouse



Type d'équipement :

- Commerces
- Enseignement
- Santé
- Services aux particuliers
- Sports, loisirs et culture
- Transports et déplacements

Nombre total d'équipements par maille :

- 100 à 167
- 50 à 100
- 10 à 50
- 1 à 10
- Aucun
- Limite communale



Sources : BPE 2017 INSEE - Admin Express IGN 2017 - Bing Aerial / Réalisation : Benjamin LARDOT - Thibaut Le Dorze / Master SIGMA / Janvier 2019



Figure 5 : Carte traitement des mailles

A l'heure actuelle, il n'y a pas de communication officielle de l'INSEE sur la mise à jour de la donnée carroyée et le maintien de sa résolution à 200 mètres. En parallèle, le projet GEOSTAT porté par l'EFGS au niveau européen promeut une trame de 1 kilomètre.

La donnée est accessible sur le site de l'INSEE :
<https://www.insee.fr/fr/statistiques/2520034>

Fichier source : ExtractHGA200mDATA

Il s'agit d'un fond de carte MapInfo au format MIF/MID qui doit être converti au format shape pour l'exploiter dans QGIS.

- Critique

Le carroyage est un mode de représentation intéressant car il constitue un découpage stable dans le temps (s'il est renouvelé) et sa résolution est plus précise que l'échelle communale.

Pour autant, l'utilisation du maillage comme échelle infra-communale comprend plusieurs limites :

- La date : La donnée carroyée ayant été réalisée en 2010, son utilisation comme indicateur de population n'est pas envisageable par rapport au nombre d'habitants issus du dernier recensement de la population en 2015.
En effet, les données des autres indicateurs étant datées de 2017 pour les équipements et 2018 pour l'activité, il s'agit de limiter l'écart entre chacun au maximum afin de limiter les biais méthodologiques pour créer l'indice de polarité.
- L'exhaustivité : La maille peut également être utilisée uniquement en tant qu'échelle de représentation. Dans cette optique, le fait que le carroyage ne couvre pas l'intégralité du territoire constitue une limite car si des équipements et des entreprises ne sont pas situés dans celui-ci, ils ne seront pas affichés. Cela pose des problèmes de représentativité considérant les pôles d'activités, les zones commerciales concentrant les équipements et entreprises sans comporter de population.
- La reproductibilité : La question de la reproductibilité se pose dans l'éventualité où l'on conserve la maille comme mode de représentation. En effet, les mailles recouvrent les zones accueillant de la population en 2010 et selon la dynamique démographique, elles pourraient désormais être vides et inversement.
Leur ré-emploi éventuel pose un problème de cohérence d'autant plus grand que l'on s'éloigne de l'année de référence.

Compte-tenu de ces éléments, le maillage n'a pas été retenu comme mode de représentation infra-communal et il a été remplacé par le point offrant une exhaustivité sur tout le territoire d'étude et plus de précision.

NB : les limites ont été échangées avec le commanditaire suite à la 1ère réunion de suivi du 23/01/19 mais la décision d'abandonner le carroyage INSEE n'est intervenue qu'à la 2nd réunion de suivi le 14/02/19 peu avant la fin du projet. Cette donnée a donc été exploitée pour des traitements intermédiaires visibles dans la suite du rapport bien qu'elle ne figure pas dans le résultat final.

3.2) Emploi

3.2.1) Nombre d'établissements employeurs du secteur privé et effectifs salariés par commune - Base SEQUOIA (ACOSS)

- Présentation

La Base SEQUOIA est fournie par ACOSS, la caisse nationale du réseau des Urssaf.

Elle propose notamment un recensement annuel du nombre d'établissements employeurs du secteur privé et des effectifs salariés par commune. La diffusion est assurée tous les ans en juin pour l'année N-1, la dernière donnée accessible étant donc pour l'année 2017.

Cette classification dans sa forme actuelle existe depuis 2007 bien que la base SEQUOIA comporte un historique depuis 1997, date à laquelle l'Acoss et les Urssaf ont débuté la centralisation des données établissements et effectifs salariés.

- Critique

Les données de la base SEQUOIA comportent de nombreux avantages comme leur exhaustivité à l'échelle nationale, la précision du décompte des effectifs salariés (pas de tranches) et le volume réduit de la table qui facilite son traitement.

Elles ont cependant été écartées car pour des raisons de secret statistique elles ne sont ni géolocalisées (x, y), ni adressées, ce qui les rend inexploitable à échelle infra-communale.

Après prise de contact auprès du correspondant Statistique Régional, on nous a indiqué qu'un projet de géolocalisation est en cours à l'Acoss et devrait aboutir à une identification des coordonnées des établissements d'ici un an. A noter que pour des raisons de secret statistique, ces données ne seront pas proposées par établissement.

3.2.2) Activité des résidents – Base IRIS (INSEE)

- Présentation

La base infracommunale IRIS « Activité des résidents » est produite par l'INSEE. Elle comporte tous les IRIS des communes ayant été découpés en IRIS et également les autres communes afin de couvrir l'ensemble du territoire.

L'activité des résidents comprend des informations détaillées sur les caractéristiques des actifs (sexe, âge, catégorie socio-professionnelle), des salariés et non-salariés (sexe et âge). On peut notamment en déduire l'effectif et le taux d'activité ou de chômage pour des résidents présents à une échelle infra-communale.

Les données sont issues du recensement de la population mis à jour chaque année par l'INSEE, dont le dernier remonte à 2015. Elles sont accessibles entre 2006 et 2015.

- Critique

Au niveau de la donnée, l'activité des résidents donne une mesure du dynamisme et de la structure économique d'une population selon son lieu d'habitation. Elle ne reflète pas l'attractivité d'un territoire et les polarités qu'elle exerce à la manière des effectifs salariés des établissements.

Concernant l'échelle, l'IRIS constitue l'élément de base des données infra-communales de l'INSEE et reste spécifique à ce fournisseur. Par conséquent, les données proposées par d'autres fournisseurs ne sont généralement pas compatibles avec l'IRIS (car la commune est souvent l'échelle de référence). Cela constitue une limite pour l'agrégation d'un indice de polarité utilisant plusieurs sources.

Par ailleurs, la précision apportée par ce niveau géographique vis-à-vis des communes ne s'applique seulement dans les communes où un découpage à l'IRIS existe à savoir les plus peuplées (au-delà de 5000 habitants).

Cette donnée n'a pas été retenue compte-tenu de ces éléments, notamment en raison du manque de précision et de compatibilité sur ce niveau d'échelle infra-communal.

3.2.3) Les stocks d'établissements et effectifs salariés par commune – Répertoire des Entreprises et des Etablissements (INSEE)

- Présentation

Le Répertoire des Entreprises et des Etablissements (REE) est fourni par l'INSEE. Il dénombre notamment les établissements et leur effectifs salariés par commune au 31 décembre de chaque année. La diffusion est assurée en fin d'année pour l'année N-2, les données étant accessibles de 2006 à 2016.

- Critique

Au niveau de la donnée, la première limite est constituée par l'utilisation de tranches pour mesurer la taille de l'effectif. Celles-ci restent relativement larges au-delà de 100 salariés avec des intervalles supérieurs à plusieurs milliers. Cette classification n'offre pas la précision satisfaisante pour constituer un indicateur d'activité.

L'échelle constitue une autre limite car comme pour la base SEQUOIA, les établissements ne sont ni géolocalisés (x, y) ni adressés et ne peuvent donc pas satisfaire une représentation infra-communale.

3.3) Démographie

Nombre d'habitants par commune - Divers (INSEE)

- Présentation

Le nombre d'habitants par commune est issu du recensement de la population de l'INSEE.

Il délivre des informations pour chaque commune sur la population, mais également sur les naissances, les décès, les résidences principales, les résidences secondaires et logements occasionnels, les caractéristiques des personnes des ménages.

Les données proviennent soit des séries historiques qui sont accessibles pour les années 1968, 1975, 1982, 1990, 1999, 2010 et 2015, soit de la base des principaux indicateurs accessibles chaque année entre 2006 et 2015.

- Critique

Bien que cette donnée soit exhaustive et précise, elle n'a pas été retenue car elle est fournie à l'échelle communale ce qui ne satisfait pas le critère de représentation infra-communal.

A noter que des données similaires correspondant au nombre d'habitants par IRIS ont été utilisées pour effectuer l'estimation de la population avec la BD TOPO.

3.4) Equipements

Couches d'équipements thématiques – Divers (Picto-Occitanie)

- Présentation

Picto-Occitanie constitue le portail interministériel cartographique des services de l'Etat pour la région Occitanie. Cette plateforme fournit notamment un catalogue de données géographiques en open-data, librement utilisable et téléchargeable.

Les fournisseurs de données correspondent aux établissements publics intervenant sur la région Occitanie, qui peuvent se constituer contributeurs ou dont les données peuvent être relayées par d'autres contributeurs.

La DREAL (Direction Régionale de l'Environnement, de l'Aménagement et du Logement) est en charge de l'administration de la plateforme et de l'animation du réseau au niveau régional. Les DDT(M) (Direction Départementales de Territoires et de la Mer) assurent en complément une animation similaire sur leurs territoires respectifs.

- Critique

Le principal intérêt est de disposer de données directement au format shapefile avec la possibilité de choisir le système de projection avant le téléchargement. Par conséquent, elles sont géolocalisées et peuvent donc être utilisées au niveau infra-communal.

De plus, les données sont proposées à l'échelle de la région l'Occitanie ou sur une échelle plus restreinte, ce qui facilite leur traitement en temps de calcul.

Cependant les données des équipements sont fournies sur certaines thématiques spécifiques (santé, éducation, sport, etc...) mais pas de façon exhaustive comme la BPE. Cela constitue une limite car toutes les thématiques ne sont pas représentées, tel que les équipements culturels et commerciaux.

Par ailleurs, les couches fournies ne comprennent pas systématiquement de métadonnées ce qui les rend difficiles à exploiter.

Compte-tenu de ces éléments, les données accessibles sur le portail Picto-Occitanie n'ont pas été retenues pour constituer l'indice de polarité.

3.5) Mobilité et transport

Pour les indicateurs de mobilité et de transport, il a été convenu avec le commanditaire suite à la seconde réunion qu'ils ne seraient pas intégrés à l'indice de polarité du fait du temps imparti limité pour recueillir les données et opérer leurs traitements.

Certaines de ces données ont néanmoins été identifiées au lancement du projet et seront détaillées ci-après.

3.5.1) Mobilités scolaires et professionnelles des individus : déplacements commune de résidence / commune de scolarisation – Fichiers détails (INSEE)

- Présentation

Ces données sont issues du recensement de la population de l'INSEE.

Les mobilités scolaires et professionnelles correspondent à des données bilocalisées domicile-travail ou domicile-études à la commune, décrivant les caractéristiques de l'individu, de son ménage et de sa résidence principale.

A partir de ces informations, il est possible de constituer un indicateur de mobilité en croisant le total des effectifs d'individus d'une commune de résidence avec la commune de travail ou d'études. Elles sont accessibles chaque année entre 2006 et 2015.

- Critique

Les mobilités scolaires et professionnelles sont trop détaillées en retraçant des parcours de mobilité par individu. Par conséquent, la table est volumineuse ce qui rend son exploitation complexe.

L'échelle de la commune constitue cependant la principale limite pour laquelle cette donnée a été écartée, ne permettant pas de répondre au critère de représentation infra-communale.

3.5.2) Flux > 100 des mobilités scolaires et professionnelles - Base flux de mobilité (INSEE)

- Présentation

Cette donnée présente les mêmes caractéristiques que les données des mobilités scolaires et professionnelles des individus, mais ne comprend que les flux de déplacements domicile-études et domicile-travail à la commune supérieurs à 100.

La base flux de mobilité est également accessible pour chaque année entre 2006 et 2015.

- Critique

Son niveau de détail convient pour constituer un indicateur de mobilité, en offrant un filtrage pour les flux marginaux < 100 individus. Cela permet également d'obtenir une table allégée aisément exploitable.

Cette donnée a cependant été écartée pour les mêmes raisons que les mobilités scolaires et professionnelles des individus avec une échelle définie à la commune.

3.5.3) Thèmes réseaux routier et ferroviaire - BD TOPO (IGN)

- Présentation

La BD TOPO est fournie par l'IGN et présente sous forme de couches vectorielles les éléments géographiques ainsi que les entités administratives et d'aménagement du territoire au niveau national.

Elle est déclinée sous différents thèmes, notamment les réseaux routier et ferroviaire qui constituent l'ensemble des infrastructures linéaires de transport pouvant être utilisés dans le cadre d'un indicateur de transport et de mobilité pour l'indice de polarité.

Les réseaux de transport de la BD TOPO peuvent ainsi être couplés à la donnée d'exploitants de transport (arrêts, desserte, fréquence...) afin de générer des isochrones.

Cette base de données est accessible sur l'espace professionnel de l'IGN et est par défaut payante, avec une mise à jour annuelle.

- Critique

La BD TOPO constitue la référence sur le plan national de par sa fréquence d'actualisation et son exhaustivité. Elle permet intrinsèquement une représentation infra-communale.

Seul son coût peut constituer une limite mais elle est mise à la disposition des services de l'Etat et des dérogations sont possibles pour y avoir accès gratuitement.

3.6) Caractérisation des surfaces qui ne sont pas dans les carreaux INSEE

Dans l'hypothèse d'un travail à partir du carroyage INSEE 2010, nous avons identifié les données qui permettraient de caractériser les surfaces exclues de ces "mailles" (donc sans données de population).

Cette caractérisation passerait par le croisement des données suivantes :

- Équipements (BPE)
- Etablissements (SIRENE)
- Population communale 2010, 2015 et son évolution (Séries historiques de l'INSEE)
- Occupation du sol (OCS) : OCS 2018 et évolution entre 2012 et 2018 (Corine Land Cover (CLC))

Ainsi, l'évolution de la population communale et celle de l'occupation du sol permettrait de contourner l'absence de données carroyée de population et de qualifier la dynamique en place sur ces territoires. Il s'agirait d'identifier les territoires situés dans des communes à forte croissance démographique et où on identifie des changements d'occupation du sol entre 2012 et 2018.

Ci-dessous, les premières étapes de ce travail :

- Préparation d'une couche commune pour étudier la population communale :

Le fichier Excel "Séries historiques" de l'INSEE est joint à une couche commune (Admin Express 2017) Nb : les séries historiques comportent les données de population de 2010 et 2015 et sont basées sur une géographie communale de 2017. Dans la couche ainsi obtenue, sont calculées les densités de population 2010 et 2015, l'évolution de la population (POP 2015 - POP 2010) et l'évolution de la densité (densité 2015 - densité 2010).

- Préparation des données d'OCS

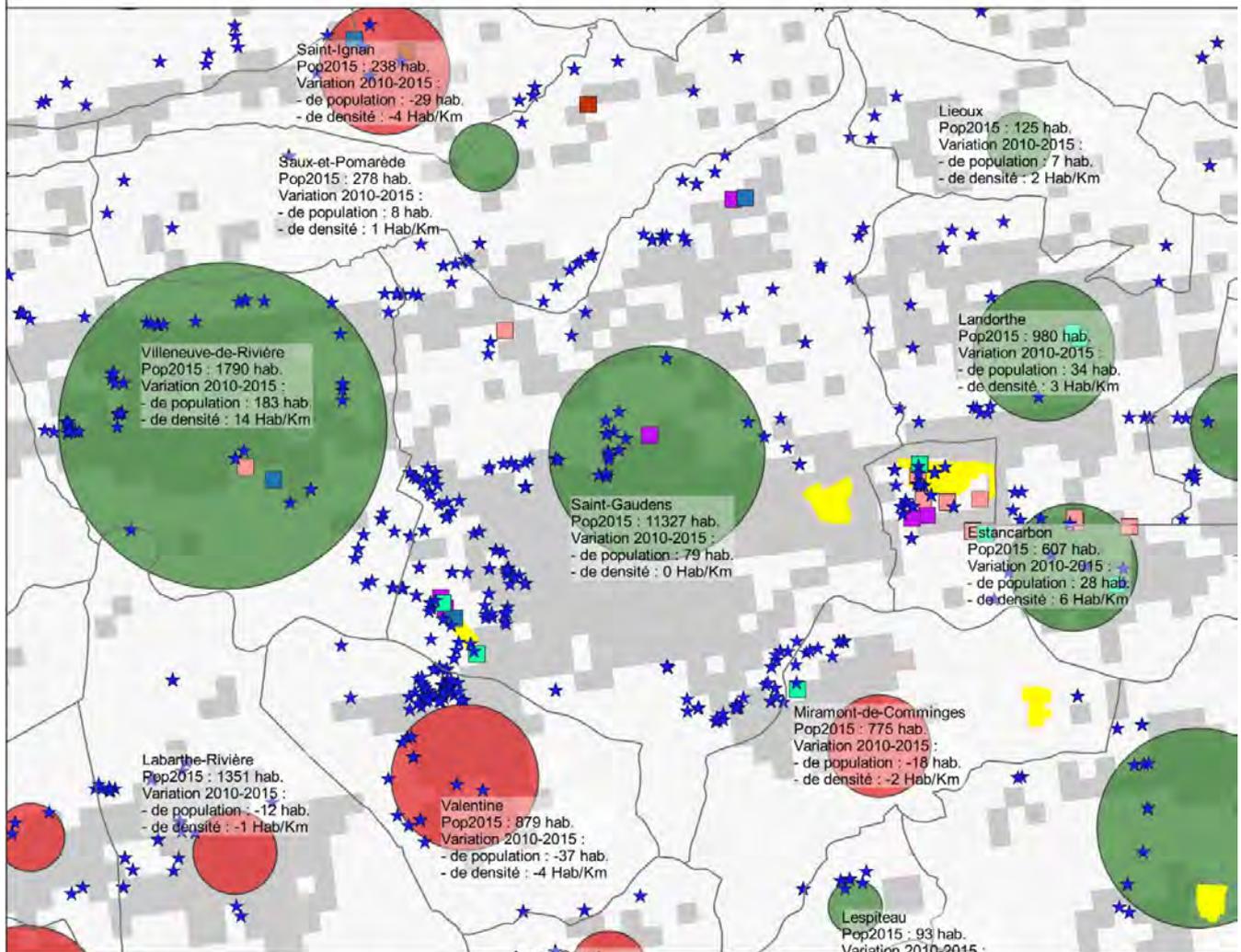
Téléchargement sur le site Copernicus des 2 couches suivantes :

- OCS 2018
- OCS 2012-2018

<https://land.copernicus.eu/pan-european/corine-land-cover/lcc-2012-2018/fetch-land-file?hash=8739958b6147e29ebe0eec340e8e1f86eb52164b>

Corine Land Cover est livrée sous forme de base de données SQLite. Ce fichier est très volumineux et nécessite donc de bonnes ressources informatiques. Il est possible de travailler avec d'autres sources : produit OSO du CESBIO ou OCS Grande Échelle de l'IGN. Corine Land Cover présente cependant l'intérêt de proposer un produit "changement 2012-2018" directement exploitable.

Données disponibles en dehors du carroyage INSEE 2010 Exemple de Saint Gaudens - Haute-Garonne (31)



Type d'équipements :

- Commerce
- Enseignement
- Santé
- Services aux particuliers
- Sports, loisirs et culture
- Transports et déplacements

- Limites communales
- Territoires couverts par la maille INSEE 2010
- Territoires ayant changé de poste dans CLC 2018 entre 2012 et 2018
- ★ Etablissements (Base SIRENE)
- Population communale en augmentation (2010-2015)
- Population communale en diminution (2010-2015)



Sources : BPE 2017 - SIRENE 2017 - Séries historiques (INSEE) / Admin Express 2017 (IGN) - CLC 2018 Réalisation : Benjamin LARDOT-Thibaut Le Dorze Janvier 2019

0 1 2 km

Figure 6 : Carte des données en dehors des mailles INSEE

4) Données utilisées

Les données utilisées pour constituer l'indice de polarité ont été retenues selon leur accessibilité, leur caractéristiques, les critères de définition et des arbitrages méthodologiques. Elles seront présentées dans cette partie, comportant également leur utilisation pratique pour le calcul de l'indice et leur critique.

4.1) Indicateur activité - Implantation d'établissements - Base SIRENE

4.1.1) Présentation

La base SIRENE des entreprises et de leurs établissements est fournie par l'INSEE et donne des informations notamment sur le nombre et le type d'établissements, ainsi que leurs effectifs salariés par tranches.

La donnée officielle est accessible sur data.gouv.fr sur la période 2017-2019 avec des mises à jour annuelles, mensuelles et quotidiennes.

Le fichier national 2017 téléchargeable sur data.gouv.fr n'est pas exploitable en l'état car trop volumineux (>8Go); en effet, il n'y a pas la possibilité de filtrer par découpage géographique avant de le récupérer.

Certains sites comme data.cquest.org et data.haute-garonne.fr proposent un filtrage par département mais sans possibilité de choisir la date de la donnée. Le site ddata.cquest.org sera privilégié car le géocodage a déjà été traité suivant une méthode exhaustive précisée sur le lien suivant :

<https://medium.com/@cq94/g%C3%A9ocodage-de-la-base-sirene-2f0e14e87a8d>

La donnée à disposition est datée du 2 décembre 2018 et a été téléchargée sur le site ci-dessous :

http://data.cquest.org/geo_sirene/last/

Fichier source : geo-sirene_31.csv

La fonction "Ajouter une couche de texte délimité (CSV)" a été utilisée pour importer la base.

La couche Shape ainsi chargée pèse 6 Go. Afin d'effectuer les traitements, la table attributaire sera allégée au maximum en supprimant les champs superflus. On obtient un nouveau Shape de 380 Mo exploitable dans QGIS.

La base SIRENE contient l'effectif salarié de chaque établissement par tranche. Les intervalles des tranches sont indiqués dans la colonne LIBTEFET.

La colonne EFETCENT donne la valeur inférieure de l'intervalle. Pour les intervalles supérieurs à 100, elle indique la valeur arrondie à la centaine inférieure.

Par exemple :

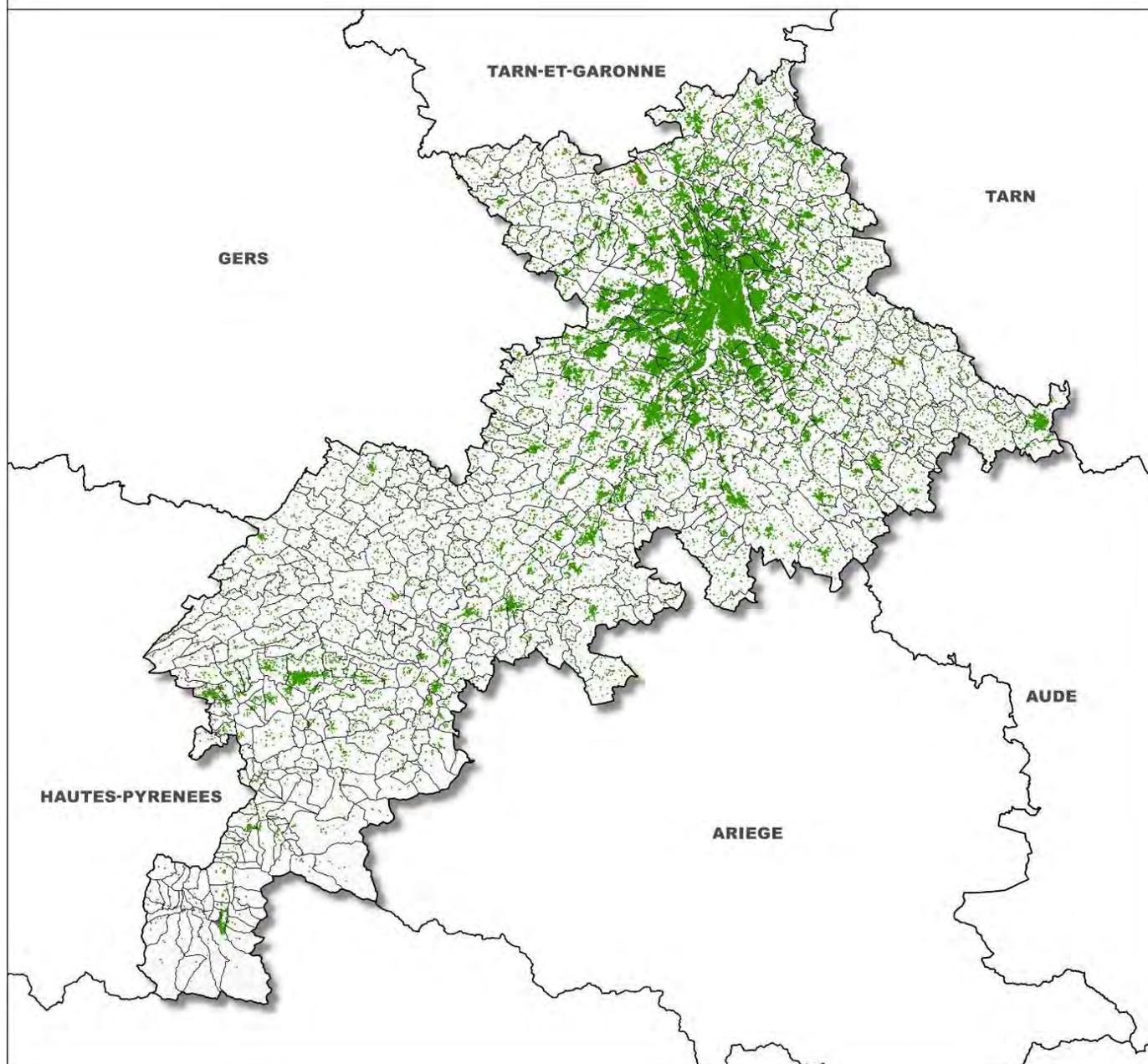
- Pour la tranche de 6 à 9 salariés, EFETCENT est égale 6
- Pour la tranche de 2 000 à 4 999 salariés : EFETCENT est égale à 2700 (dans ce cas l'effectif est situé entre 2700 et 2800 salariés)

Cette colonne sera utilisée pour obtenir une estimation de l'effectif salarié.

A noter que :

- La modalité 0 signifie que l'établissement n'a pas d'effectif au 31/12 mais a employé des salariés au cours de l'année de référence.
- La modalité NN correspond aux unités non employeuses, qui n'ont pas de salariés au cours de l'année de référence et pas d'effectif au 31/12. Cette différenciation n'est pas utile pour l'étude et on leur attribuera la valeur 0.

Présentation des données sources
Etablissements recensés dans la base SIRENE 2018 de l'INSEE
Département de Haute-Garonne



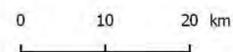
Fichier.csv géolocalisé

218 689 établissements :

-72 % sont géolocalisés dans le carroyage INSEE 2010 (156 904 établissements)

-28 % sont géolocalisés en dehors du carroyage INSEE 2010 (61 781 établissements)

-Seuls 18 % des enregistrements sont renseignés avec au moins un salariés.



Sources : SIRENE 2018 INSEE - Admin Express IGN 2017 / Réalisation : Benjamin LARDOT - Thibaut Le Dorze / Master SIGMA / Janvier 2019

Figure 7 : Carte de présentation des données SIRENE

4.1.2) Application

- On crée une couche Shape à partir du fichier source geo-sirene_31.csv
- On supprime dans QGIS les champs superflus pour alléger la table attributaire et la rendre exploitable
- On effectue une intersection entre le fichier shape des établissements et le département de Haute-Garonne pour ne retenir que les établissements situés dans le périmètre d'étude (car certains établissements, mal géocodés apparaissent en dehors du département)
- On supprime les établissements ayant un géocodage inférieur à 0,6 en utilisant le champ geo_score
- On remplace les valeurs 'NN' du champ EFETCENT par 0
- On sélectionne les établissements en utilisant l'option intersekte de la sélection par localisation à partir des couches SCOT_TOULOUSE et SCOT_HORS_TOULOUSE et on enregistre la sélection

Cf. Annexe pages 57-58 : Diagramme de flow - Etablissements

4.1.3) Critique

La base SIRENE a été privilégiée pour constituer l'indicateur d'activité car il s'agit de la seule donnée relative aux établissements géolocalisée, ce qui permet d'assurer une représentation au niveau infra-communal.

Le géocodage possède une qualité variable selon les établissements qui est répertoriée dans le champ geo-score. Bien que les établissements ayant une valeur de géocodage inférieure à 0,6 aient été retirés, il subsiste néanmoins des poches d'accumulation d'établissements en un même point possédant un geo-score important. Celles-ci peuvent correspondre à la réalité comme à un géocodage imprécis, ce biais méthodologique devant être intégré dans l'interprétation des résultats.

La base SIRENE comporte également des limites concernant les effectifs salariés :

- Elles sont égales à NN ou à 0 pour 82% des établissements, ce qui tend à concentrer le poids de cette variable sur une minorité des établissements.
- Les valeurs des effectifs salariés des établissements étant renseignés par tranches arrondies à la centaine, cela n'offre qu'une estimation et ne peut être interprété comme un dénombrement.

Compte-tenu de ces éléments, les effectifs salariés ont été écartés de l'étude et seul le nombre d'établissements a été conservé comme variable pour constituer l'indicateur d'activité.

4.2) Indicateur démographique - Estimation de la population

4.2.1) Présentation

La méthode consiste à répartir la population des IRIS dans les bâtiments localisés dans chaque IRIS en fonction de leur surface habitable.

NB : Après avoir estimé la population pour chaque bâtiment, il est possible d'attribuer aux mailles le total de la population des bâtiments qu'elles contiennent. Cette option n'a pas été utilisée dans le cadre de l'indice de polarité mais la méthode a été explicitée dans la partie Application.

Cette méthode est adaptée de la méthode développée par le Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA), consultable sur le lien suivant (4.5.3 Méthodologie BD TOPO p57-59) :

https://www.lcsqa.org/system/files/drc-15-144366-01026a_methodologiepopulation_vf.pdf

NB : Ce document indique que la cette méthodologie est pour partie documentée dans des fiches CERTU (ex-CEREMA) élaborées en 2005 et 2006 mais ces dernières ne sont pas accessibles en ligne.

L'estimation de la population nécessite donc trois données (Contours IRIS, Base IRIS et BD TOPO) :

- Contours IRIS

Les contours des IRIS de l'INSEE correspond à un fond numérisé des îlots pour les besoins des recensements de la population 2011 sur l'ensemble des communes de plus de 10 000 habitants et la plupart des communes de 5 000 à 10 000 habitants.

La base Contours IRIS Édition 2017 est disponible sur le lien suivant (données France entière version 2.1) :

<http://professionnels.ign.fr/contoursiris>

On sélectionne ensuite uniquement les IRIS situés dans le département de Haute-Garonne.

- Base IRIS

La base infra-communale IRIS issue du Recensement de la Population 2015 :

<https://www.insee.fr/fr/statistiques/3627376>

On effectue une jointure entre la base et les contours IRIS pour récupérer les valeurs de population par IRIS en 2015.

- BD TOPO

La BD TOPO est fournie par l'IGN et correspond à une base de données vectorielles de l'ensemble des entités géographiques et administratives du territoire national. Elle est accessible sur l'espace professionnel de l'IGN et est par défaut payante, avec une mise à jour annuelle. La BD TOPO 2015 est mise à disposition par la DDT Haute-Garonne pour l'usage spécifique du projet. Cette version est utilisée pour disposer de la même année de référence que la base IRIS (la plus récente étant issue du recensement de la population de 2015).

Le fichier source au format MapInfo a été converti en shape pour l'exploiter dans QGIS.

Présentation des données sources : Population 2015



La population de chaque bâtiments a été recalculée à partir de la BD TOPO 2015 (IGN) et de la base IRIS 2015 (INSEE). Cette population est ensuite injectée dans le carroyage INSEE 2010



- Bâtiments résidentiels
- Carroyage INSEE 2010

Sources : SIREN 2018 INSEE - Admin Express IGN 2017 / Réalisation : Benjamin LARDOT - Thibaut Le Dorze / Master SIGMA / Janvier 2019

0 25 50 m

Figure 8 : Carte de présentation des données démographiques

4.2.2) Application

a) Sélection du bâti résidentiel

- On sélectionne la couche BATI_INDIFFERENCIE : cette classe regroupe tous les bâtiments de la BD TOPO dont le caractère n'a pas pu être décrit (hors BATI_INDUSTRIEL et BATI_REMARQUABLE) dont les logements, certains commerces, de l'enseignement...
- On supprime les bâtiments qui touchent un Point d'Activité d'Intérêt (toutes les couches PAI ponctuelles de la BD TOPO), à l'exception des champs "Châteaux" et "Lieu-dit habité" pour le PAI_Zone_Habitat qui sont les seuls correspondant à une zone d'habitation
- On supprime également les bâtiments qui touchent une surface d'activité (couche SURFACE_ACTIVITE du thème Zone d'activités de la BD Topo)
- On supprime les bâtiments de surface faible inférieure à 40 m² car supposés être des bâtiments non appropriés pour un usage d'habitation

b) Estimation de la surface d'habitation

- On supprime les bâtiments de hauteur inférieure à 3m, car considérés comme étant des garages ou des dépendances
- On estime le nombre d'étages par bâtiment selon sa hauteur : un seul niveau pour les bâtiments jusqu'à 5 mètres puis un niveau supplémentaire tous les 3m
 - De 3 à 5m : 1 étage
 - De 6 à 8m : 2 étages
 - Etc...
- On réalise une intersection entre le bâti et l'IRIS pour et recalcule les surfaces des polygones obtenus
- On calcule ensuite la surface habitable : surface au sol * nombre d'étages

c) Estimation de la population par bâtiment

- On sélectionne la surface habitable totale des bâtiments par IRIS en utilisant le gestionnaire de base de données

Script SQL :

```
select BATI_SURF_HAB.*, sum(BATI_SURF_HAB.SURF_HAB) as SURF_HAB_I,  
BATI_SURF_HAB.geometry  
from BATI_SURF_HAB  
group by BATI_SURF_HAB.CODE_IRIS
```

- On effectue une jointure entre la couche obtenue par la requête et la couche des bâtiments pour récupérer le champ comprenant la surface habitable totale des bâtiments par IRIS
- On calcule ensuite le poids normalisé de chaque bâtiment à l'échelle de l'IRIS (surface du bâtiment/surface totale des bâtiments de l'IRIS)
- On multiplie ensuite la population totale de l'IRIS par le poids normalisé du bâtiment à l'échelle de l'IRIS pour obtenir la population par bâtiment

d) Sélection de l'aire d'étude

- On sélectionne les bâtiments en utilisant l'option intersecte de la sélection par localisation à partir des couches SCOT_TOULOUSE et SCOT_HORS_TOULOUSE et on enregistre la sélection

d) (Optionnel) Distribution de la population des bâtiments par maille et sélection de l'aire d'étude

- On réalise une intersection entre les polygones de chaque bâtiment et les mailles et on recalcule la surface des polygones obtenus
- On supprime les polygones dont la surface est inférieure à 0.01 car leur poids est marginal
- On calcule le ratio entre la surface des polygones intersectés avec les mailles et celle des polygones initiaux
- On multiplie ensuite la population du bâtiment d'origine par le ratio pour obtenir la population par bâtiment
- On sélectionne la population totale des bâtiments contenus dans chaque maille en utilisant le gestionnaire de base de données

Script SQL :

```
select BATI_POP_MAILLES.*, sum(BATI_POP_MAILLES.POP_MAIL) as  
POP_2015, BATI_POP_MAILLES.geometry  
from BATI_POP_MAILLES  
group by BATI_POP_MAILLES.idINSPIRE
```

- On effectue une jointure entre la couche obtenue par la requête et la couche des mailles pour récupérer le champ créé
- On convertit les valeurs nulles du champ POP_2015 en 0 (cela correspond à l'absence de bâtiment dans certaines mailles)
- On cartographie à partir de ce champ la population totale des bâtiments par maille
- On sélectionne les bâtiments en utilisant l'option intersecte de la sélection par localisation à partir des couches SCOT_TOULOUSE et SCOT_HORS_TOULOUSE et on enregistre la sélection

Cf. Annexe pages 59-62 : Diagramme de flow – Population par bâtiment

L'alternative sur la distribution de la population des bâtiments par maille est également présentée :

Cf. Annexe pages 63-67 : Diagramme de flow - Population par maille

4.2.3) Critique

L'avantage de cette méthode est de pouvoir obtenir des valeurs de population à partir d'une donnée existante en modulant l'échelle de représentation : on peut ainsi passer d'une démographie à l'IRIS à une démographie au bâti puis à la maille.

Cependant cela reste une estimation qui peut comprendre des biais, principalement dans la sélection du bâti résidentiel depuis la couche BATI_INDIFERENCIE. La méthode LCSQA préconise en effet d'effectuer une vérification manuelle des bâtiments à exclure touchant un PAI ou une surface d'activité, car un filtrage automatique peut comporter des erreurs. Il s'agit là de valider chaque bâtiment à supprimer en les identifiant via l'imagerie satellite, navigation Street View ou visite terrain afin de vérifier qu'ils ne soient pas des habitations.

Ce procédé n'a pas été effectué compte-tenu de la quantité d'objets traités à l'échelle du département, ce qui induit intrinsèquement des erreurs.

De même les autres filtres de sélection du bâti résidentiel (superficie, hauteur etc....) et l'attribution du nombre d'étages sont schématiques et n'intègrent pas toutes les variables présentes en situation réelle (ex : un bâti ancien possède des étages plus hauts que le bâti récent ce qui peut fausser le calcul de la surface habitable).

La méthode LCSQA de répartition de la population des IRIS dans les bâtiments doit ainsi être nuancée au regard de ces éléments et bien considérée en tant qu'estimation.

L'estimation de la population ayant été faite pour chaque bâtiment répartis ensuite par SCOT avec une sélection par localisation, les bâtiments dont l'emprise touche deux SCOT qui seront rattachés à un seul SCOT arbitrairement. Cela concerne uniquement 78 bâtiments sur les 201965 bâtiments retenus sur le département. Il existe donc un biais méthodologique même s'il reste marginal.

Cette méthode de simplification a été retenue par manque de temps compte-tenu des changements de zone d'étude intervenus tardivement dans l'étude. Idéalement, la répartition de la population estimée d'un bâtiment sécant à deux SCOT doit être effectuée selon le ratio de sa surface intersectée.

4.3) Indicateur d'équipement - Implantation d'équipements

4.3.1) Présentation

La base permanente des équipements (BPE) est mise à disposition chaque année par l'Insee. Élaborée à partir de différentes sources administratives, cette base répertorie un large éventail d'équipements et de services rendus à la population. Elle recouvre les domaines des services, marchands ou non, des commerces, de la santé et de l'action sociale, de l'enseignement, du sport, du loisir, du tourisme, des transports et de la culture.

Nous utiliserons la dernière version (2017), seule édition accessible.

La base est proposée sous forme d'un fichier .dbf de 49 627 enregistrements dont la qualité de géolocalisation est qualifiée dans un champ "qualité_xy" avec 4 modalités :

- "Non géolocalisés" : 11 725 équipements
- Géolocalisation « mauvaise » : 2 501 équipements
- Géolocalisation « acceptable » : 1 441 équipements
- Géolocalisation « bonne » : 33 960 équipements

NB : les équipements non géolocalisés concernent l'artisanat du bâtiment et certains services automobiles (contrôle technique, location auto-utilitaire et école de conduite).

Seuls les équipements avec une géolocalisation « bonne » ou « acceptable » sont retenus, soit 35 401 entités représentant environ 70% de la BPE 2017.

L'INSEE fournit également un fichier .xls « gammes d'équipements ». Ce fichier permet de regrouper les enregistrements en "gammes" et a pour objectif de réunir des équipements ayant des logiques d'implantation similaires, c'est-à-dire des équipements étant fréquemment situés à proximité. Le fichier « gammes d'équipements » permet ces regroupements d'équipements par jointure sur la base BPE.

Pour la BPE 2017, 110 équipements sont retenus pour constituer trois niveaux de gammes :

- Gamme supérieure qui regroupe 47 types d'équipements : 2 857 équipements dans la BPE 2017
- Gamme intermédiaire qui regroupe 36 types d'équipements 9 127 équipements dans la BPE 2017
- Gamme de proximité qui regroupe 27 types d'équipements 22 267 équipements dans la BPE 2017

NB : lorsque ce fichier de regroupement est joint à la base BPE, 1 162 équipements disparaissent sur les 35 401 équipements géolocalisés faute de champ commun. Sur ces 1 162 enregistrements, 700 correspondent à des équipements à caractère touristique, des équipements de loisirs ainsi qu'à certains équipements sportifs. Ces derniers fournissent leurs services davantage à une clientèle de passage qu'à la population résidente. La perte des équipements liée à la jointure concerne donc 462 équipements utiles à la base.

Le fichier "gamme d'équipements" permet également de qualifier les équipements selon leur type (commerce, enseignement, culture...)

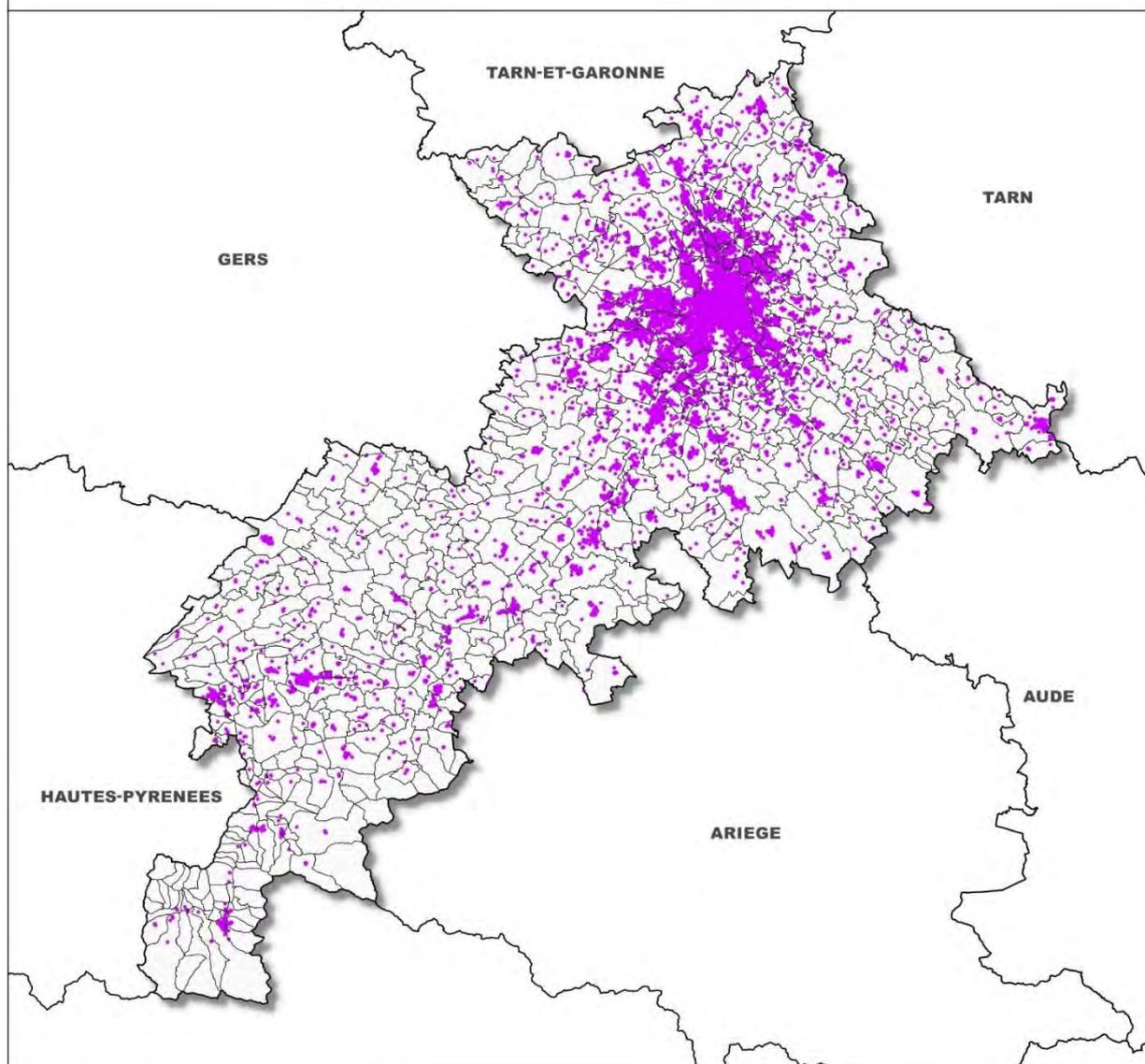
Finalement, 34 239 équipements sur les 49 627 contenus dans la base sont exploitables (68.9%) dont :

- 6 133 commerces
- 1 144 équipements d'enseignement
- 11 394 équipements de santé
- 11 995 équipements de services aux particuliers
- 2 789 équipements de sport, loisir et culture
- 784 équipements de transport

Soit :

- 2 857 équipements dans la gamme supérieure
- 9 127 équipements dans la gamme intermédiaire
- 22 267 équipements dans la gamme de proximité

Présentation des données sources
Équipements recensés dans la Base Permanente des Equipements (BPE) 2017
de l'INSEE - Département de la Haute-Garonne

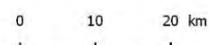


Fichier .dbf géolocalisé (49 627 Equipements) :

30 % non géolocalisés (14 926)
 65 % géolocalisés dans le carroyage INSEE 2010 (32 119)
 5 % géolocalisés en dehors du carroyage INSEE 2010 (2 582)

Parmi les équipements géolocalisés:

- | | |
|--|---|
| -6133 commerces | -Gamme proximité : 22 267 équipements |
| -1144 équipements d'enseignement | -Gamme intermédiaire : 9127 équipements |
| -11394 équipements de santé | -Gamme supérieure : 2857 équipements |
| -11 995 équipements de services aux particuliers | |
| -2789 équipements de sport, loisir et culture | |
| -784 équipements de transport | |



Sources : BPE 2017 INSEE - Admin Express IGN 2017 / Réalisation : Benjamin LARDOT - Thibaut Le Dorze / Master SIGMA / Janvier 2019

Figure 9 : Carte de présentation des données d'équipement

4.3.2) Application

Nous avons subdivisé le contenu de la base de différentes façons afin d'optimiser son exploitation et de mener à bien divers tests.

Ainsi, dans notre cas, il y a deux façons d'utiliser les données BPE :

- Soit en utilisant les données brutes (géolocalisées au point)
- Soit en regroupant les équipements contenus dans une même maille INSEE 2010 ("géolocalisation" par maille). Nous avons expérimenté les deux façons de faire.

Ensuite, nous avons créé des regroupements par type d'équipement (santé, commerce, enseignement, transport, service aux particuliers et sport-loisir-culture) et par gamme (proximité, intermédiaire et supérieure).

Au final, ce sont donc 20 couches shapefile ponctuelles qui sont créées :

Total :

- Équipements total
- Nombre total d'équipements par maille

Par type :

- Équipements de santé
- Équipements de commerce
- Équipements d'enseignement
- Équipements de transport
- Équipements de sport, loisir et culture
- Équipements de service aux particuliers

- Nombre d'équipements de santé par maille
- Nombre d'équipements de commerce par maille
- Nombre d'équipements d'enseignement par maille
- Nombre d'équipements de transport par maille
- Nombre d'équipements de sport, loisir, culture par maille
- Nombre d'équipements de services aux particuliers par maille

Par Gamme :

- Équipements de proximité
- Équipements intermédiaires
- Équipements supérieurs

- Nombre d'équipements supérieurs par maille
- Nombre d'équipements intermédiaires par maille
- Nombre d'équipements de proximité par maille

Chacune de ces couches fera l'objet de tests afin d'estimer la pertinence de les utiliser dans l'élaboration d'un indice de polarité.

PS : Qualification des équipements situés en dehors des mailles INSEE 2010 (2 582 équipements) :

- 822 commerces
- 84 enseignement
- 313 santé
- 755 services aux particuliers
- 488 sport, loisir, culture
- 30 transport et déplacement
- 90 non renseignés

Cf. Annexe page 68 : Diagramme de flow – Equipements

4.3.3) Critique

La BPE étant constituée à partir de différentes sources, la qualité de la géolocalisation est hétérogène. Ce biais pourrait être limité en ne travaillant qu'à une échelle communale mais la DDT souhaite avant tout une illustration infra-communale du phénomène.

Par ailleurs, dans notre recherche bibliographique, nous avons constaté que les études sur les polarités de services ne s'appuient que rarement sur la BPE. Le plus souvent, ce sont des enquêtes terrain qui permettent d'obtenir des données fiables sur les équipements.

Le fichier "gamme d'équipement" est également problématique ; il n'a visiblement pas été correctement mis en cohérence avec la base. Lors de la jointure, des équipements disparaissent. De plus, certains choix sont discutables. Par exemple, les écoles maternelles sont considérées comme des équipements intermédiaires quand les écoles élémentaires sont classées dans les équipements de proximité.

La BPE reste néanmoins une donnée intéressante dans une perspective de reproduction de notre travail. Ainsi, même si la qualité de la donnée est discutable, sa mise à jour annuelle rend les comparaisons possibles d'année en année.

5) Paramètres du plugin « Carte de chaleur » de QGIS

C'est avec ce plugin que nous avons traité toutes les couches ponctuelles préparées en amont. Cet outil permet de visualiser la concentration d'un phénomène ponctuel. Il génère des cartes de chaleur sous forme de raster à partir d'une couche de points.

Cela permet d'identifier facilement des "points chauds", de visualiser la concentration d'un phénomène ponctuel.

Le plugin "carte de chaleur" proposé dans QGIS utilise la méthode des noyaux pour créer une carte de chaleur à partir d'une couche vecteur de points. La densité de noyaux est calculée en fonction du nombre de points présents dans un périmètre fixé autour de chaque point ; plus le nombre est important plus la valeur est grande.

La méthode des noyaux est donc une méthode d'interpolation spatiale permettant d'estimer des **surfaces de densité** à partir d'un nuage de points, de mettre en évidence des différentiels de concentration, et donc des polarités, à partir de leurs positions dans l'espace.

Le paramétrage du plugin doit être étudié avec soins : les résultats et leur interprétation en dépendent. Nous avons donc effectué de nombreux tests afin de bien comprendre le fonctionnement de l'outil et de mesurer l'influence de chacun des paramètres

Les différents paramètres sont :

5.1) Le Rayon

Le rayon d'influence est un paramètre déterminant. Il est utilisé pour définir le rayon de recherche de la carte de chaleur (en mètre ici) autour de chaque point. Il permet donc d'une certaine façon de pondérer le poids d'une donnée. Le rayon définit la distance autour d'un point au-delà de laquelle son influence sera nulle. Les valeurs les plus grandes donnent un plus grand lissage (car plus de points se chevauchent), mais des valeurs plus petites génèrent plus de détails et de variations en densité de points (moins de chevauchement). Il aurait été intéressant d'utiliser un rayon "variable" (plus grand dans les zones avec peu de points et plus faible dans les zones plus denses en points) mais cette fonction n'existe pas sous QGIS.

Plusieurs rayons ont été testés : (50, 100, 200, 500, 1000, 2000 et 5000 m).

Le rayon est à choisir en fonction des éléments ponctuels à représenter (leurs répartitions spatiales, leurs densités, leur importance relative)

Les polarités les plus petites doivent apparaître ; à l'inverse les polarités ne doivent pas être trop diffuses. A l'intérieur des polarités, les différences de valeurs doivent rester perceptibles.

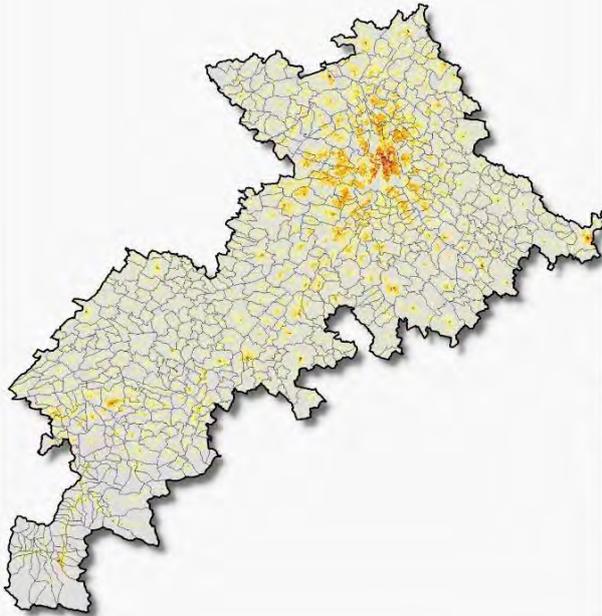
Pour la population, nous avons retenu un rayon de 200 m. En effet, chaque "point" de population étant en réalité le centroïde d'un bâtiment, il faut utiliser un rayon assez faible afin de ne pas trop lisser le résultat.

Pour les données SIRENE, nous avons retenu un rayon de 500 m. La base contenant plus de 200 000 enregistrements sur l'ensemble du département, il s'avère après tests qu'un diamètre supérieur lisse trop le résultat.

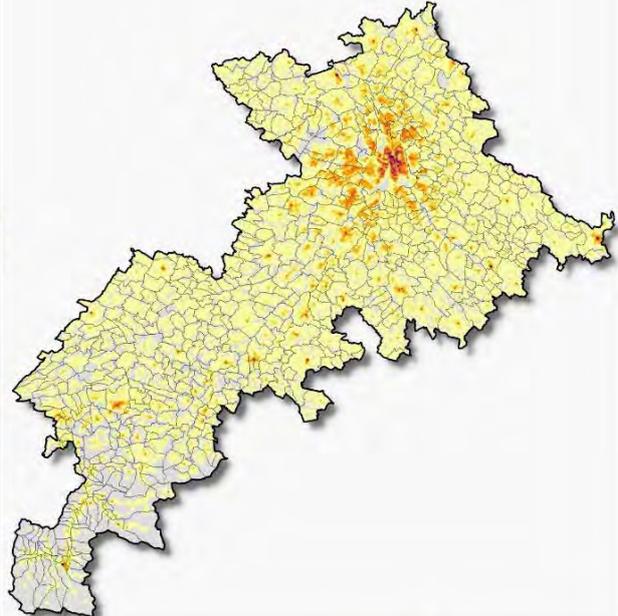
Concernant les équipements, nous avons augmenté le rayon de recherche de la carte de chaleur avec la rareté des équipements : 500 mètres pour les équipements de proximité, 1000 pour les équipements intermédiaires et 5000 mètres pour les équipements supérieurs.

**Influence du rayon de recherche dans l'élaboration d'une carte de chaleur :
Exemple de la population en Haute-Garonne (31)**

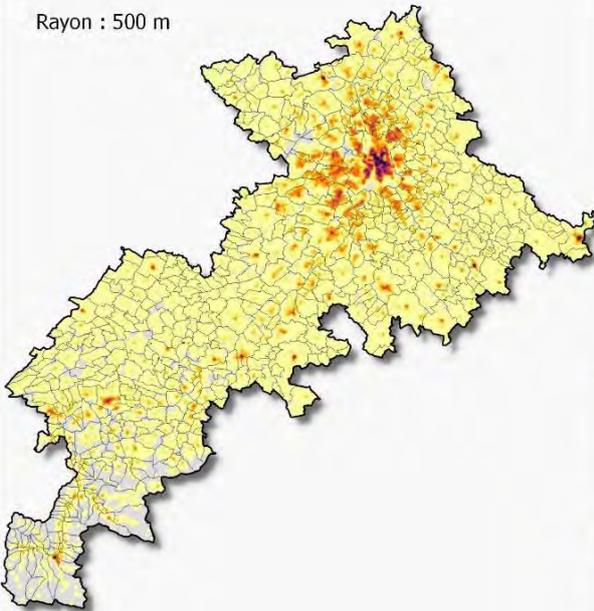
Rayon : 100 m



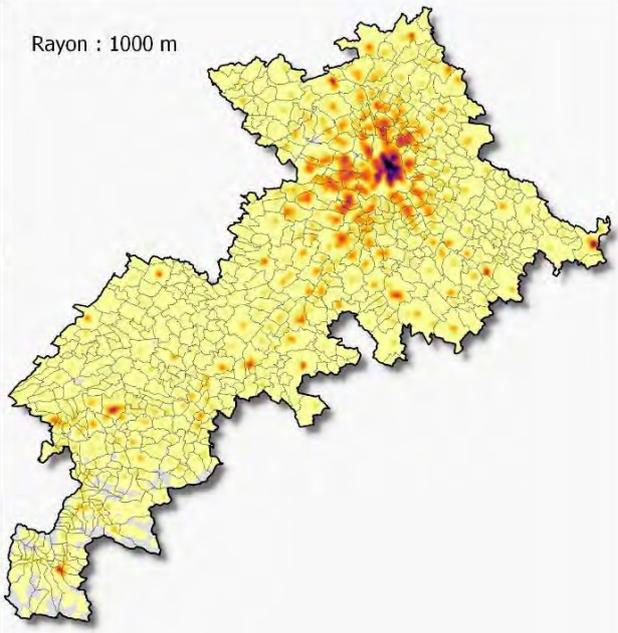
Rayon : 300 m



Rayon : 500 m



Rayon : 1000 m



Carte élaborée avec la population calculée de chaque bâtiment résidentiel placée au centroïde des bâtiments

Source : BD TOPO IGN 2015, BD IRIS INSEE 2015, ADMIN EXPRESS IGN 2017 / Réalisation : Benjamin LARDOT Thibaut Le Dorze / Master SIGMA / Janvier 2019

0 10 20 km

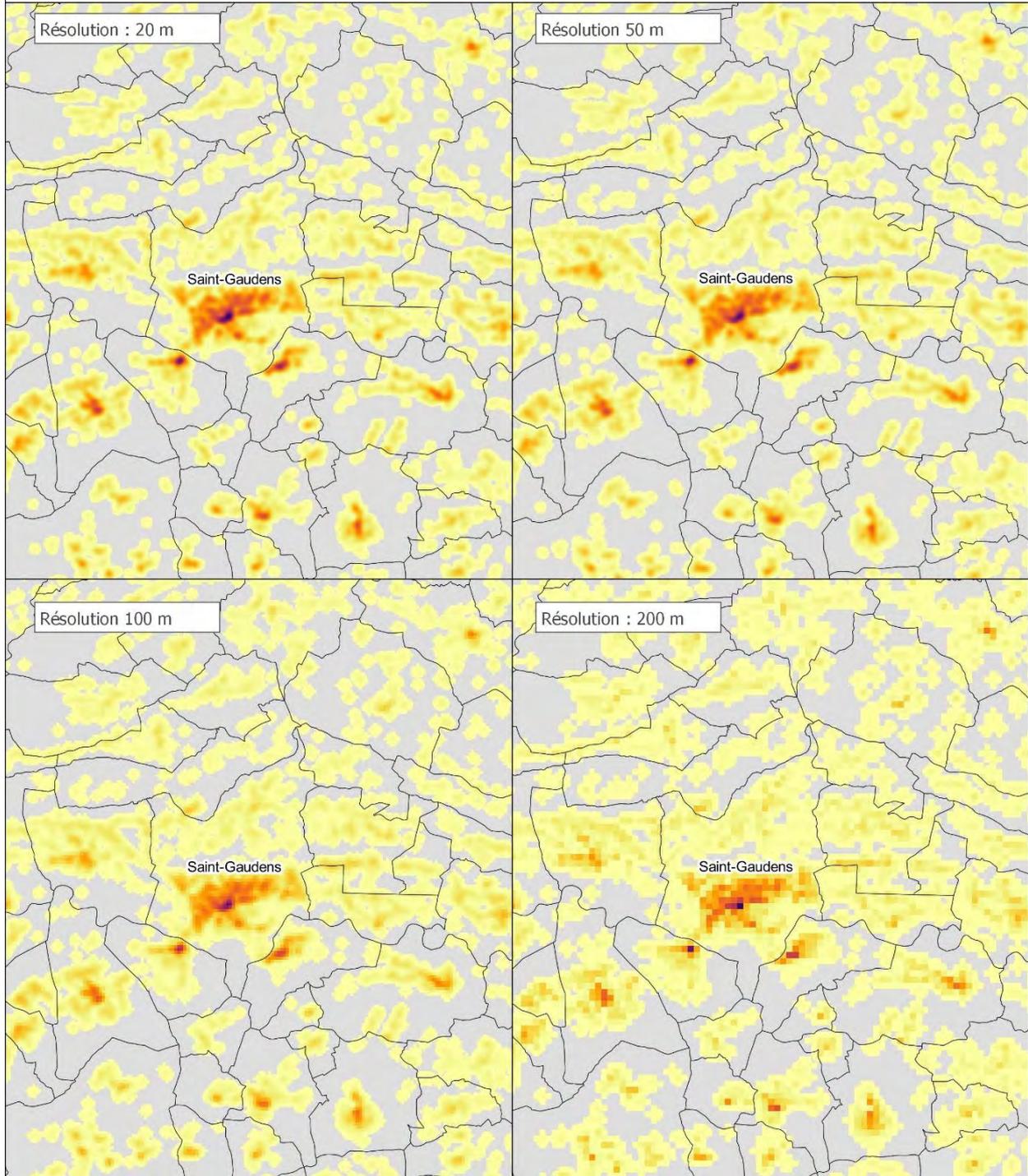


Figure 10 : Carte de test du rayon

5.2) Taille de la cellule

Il s'agit de la résolution du raster de sortie. Une grille fine conduira à opérer un lissage faible. A l'inverse, une cellule de grande dimension se traduira par un lissage important. La taille de cellule correspond finalement la taille de pixel de l'analyse finale souhaitée. Nous utiliserons une cellule de 20 mètres (donc assez fine) car chaque raster de chaleur généré sera par la suite aligné aux autres puis (peut-être) ré-aggloméré en maille de 200 mètres générant ainsi autant d'effets de lissage, d'imprécisions supplémentaires. Nous choisissons donc une résolution fine.

Influence de la résolution du raster de chaleur dans l'élaboration d'une carte de chaleur
Exemple de la population, région de Saint-Gaudens en Haute-Garonne (31)



Source : BD TOPO IGN 2015, BD IRIS INSEE 2015, ADMIN EXPRESS IGN 2017 / Réalisation : Benjamin LARDOT Thibaut Le Dorze / Master SIGMA / Janvier 2019

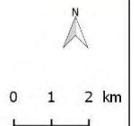


Figure 11 : Carte de test de la résolution

5.3) Forme du noyau

La forme du noyau contrôle le taux à laquelle l'influence d'un point diminue à mesure que la distance au point augmente (c'est le lissage spatial). Différents noyaux diminuent à des taux propres (selon différentes fonctions) : Triweight, Epanechnikov, Quartique, Triangulaire ou Uniforme.

Le noyau uniforme ne permet pas de faire varier la densité selon la distance (la valeur de densité est la même en tous points autour de l'équipement).

Pour le noyau triangulaire, la densité diminue de façon linéaire depuis le point vers l'extérieur du rayon de recherche.

Le noyau quartique donne un poids plus grand aux points proches qu'aux points éloignés, mais la décroissance est graduelle. Il en est de même pour les noyaux Triweight et Epanechnikov mais avec d'autres fonctions (Cf. schémas ci-dessous)

Il est par ailleurs admis que le choix de la fonction de densité est bien moins important que celui du rayon d'influence.

Nous utiliserons finalement le noyau Quartique qui est celui employé par défaut par QGIS et la seule option proposée par ArcGIS.

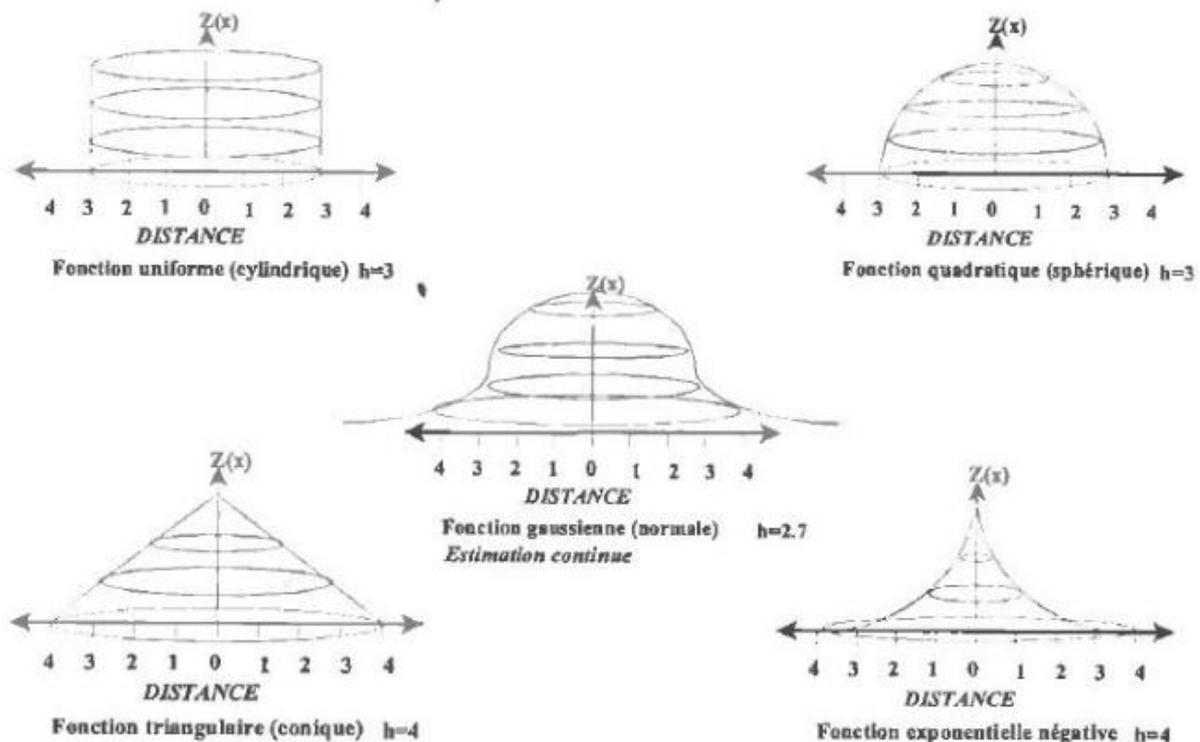
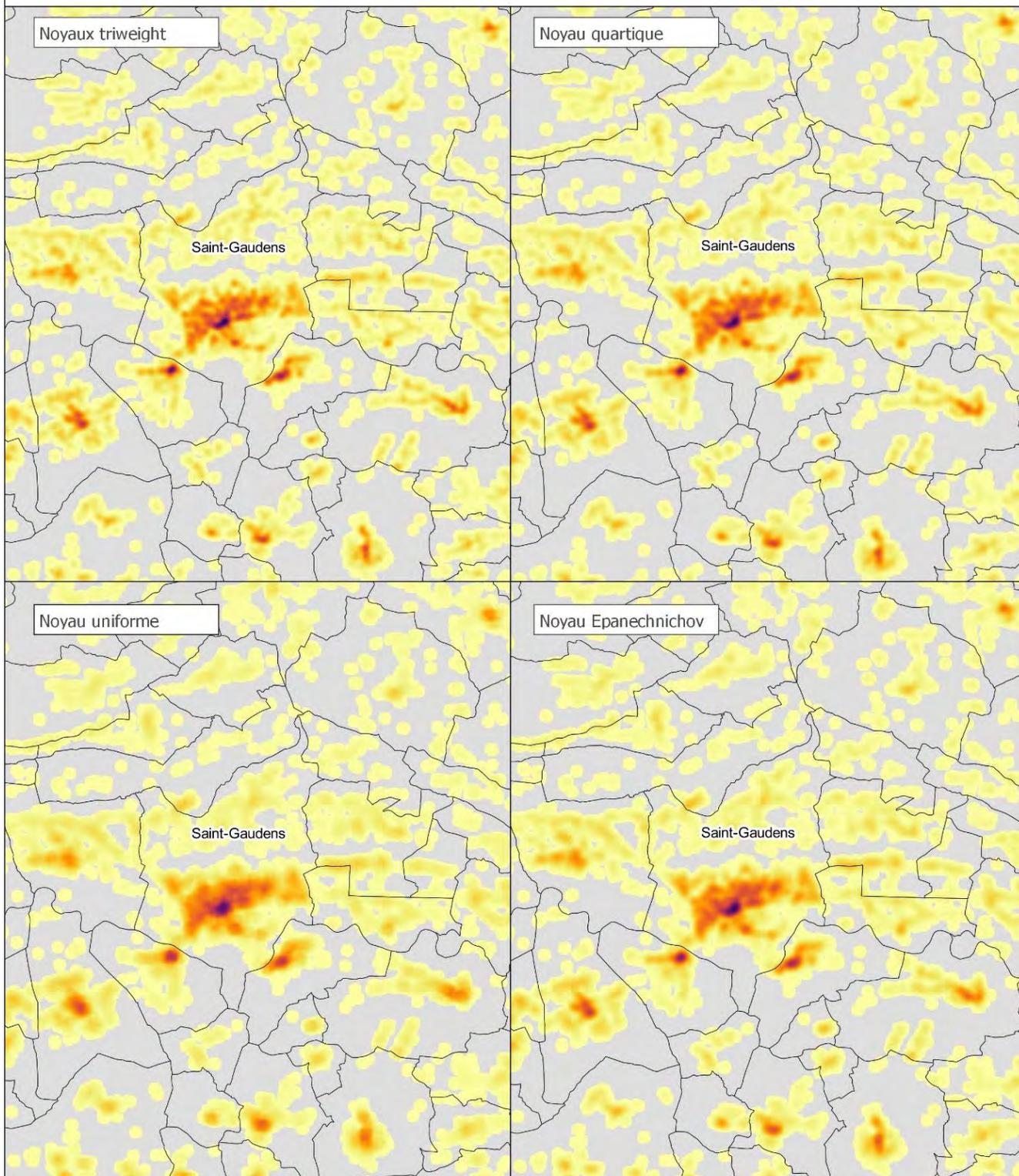


Figure 12 Illustration des fonctions utilisées pour calculer une densité de noyau

Influence du noyaux utilisé dans l'élaboration d'une carte de chaleur :
Exemple de la population, région de Saint-Gaudens en Haute-Garonne (31)



Source : BD TOPO IGN 2015, BD IRIS INSEE 2015, ADMIN EXPRESS IGN 2017 / Réalisation : Benjamin LARDOT Thibaut Le Dorze /Master SIGMA / Janvier 2019

0 10 20 km

Figure 13 : Carte de test du type de noyau

5.4) Pondération :

L'importance de certaines entités ponctuelles peut être appuyé par une pondération à partir d'un attribut de la couche de point. C'est ce que nous avons fait avec la donnée sur les établissements (SIRENE) : cette donnée est pondérée par son champ "géo score" afin de donner un poids plus important aux établissements les mieux géolocalisés (proches de 1). Les autres données utilisées ne sont pas pondérées à ce niveau.

6) Elaboration de l'indice de polarité

Les établissements, la population et les équipements ne sont pas répartis uniformément sur le territoire départemental. Ainsi la métropole de Toulouse "aspire" quasiment les $\frac{2}{3}$ de chaque donnée. Les fortes concentrations en région toulousaine ont pour conséquence de minimiser les polarités plus rurales. Pour pallier à ce problème, nous avons divisé notre travail en deux lots :

Un premier traitement à l'échelle du SCoT Grande Agglomération Toulousaine puis un second traitement à l'échelle départementale (SCoT toulousain exclu).

Ci-dessous, en résumé, les différentes étapes conduisant à l'élaboration de l'indice de polarité.

6.1) Disposer de données sources ponctuelles

(Centroïdes des mailles ou points)

Cf. Diagrammes de flow Equipement, Etablissements, Population par bâtiment et Population par maille.

6.2) Ajout de 4 points à chaque couche.

Cela permet de fixer l'étendue des traitements ultérieurs sur chaque couche.

6.3) Traitement de chaque nuage de points par le plugin "carte de chaleur".

Utilisation des paramètres adéquats (rayon, taille de cellule, type de noyaux, pondération éventuelle).

Ce traitement génère des rasters de chaleur élémentaires pour chaque donnée. Afin d'assurer la cohérence de tous les rasters, la taille de cellule de sortie est toujours fixée à 20 mètres ; le rayon, lui, varie selon la donnée traitée. Nous n'avons utilisé que le noyaux quartique et seule la couche SIRENE est pondérée par son champ "geo score".

6.4) Alignement des rasters

Réalisé avec l'outil "Aligner des rasters". Ce traitement est nécessaire pour correctement combiner les différentes images par la suite.

6.5) Suppression des valeurs -9999

L'outil précédent, "Aligner des rasters", génère des valeurs à -9999 pour les pixels sans données. Ces valeurs doivent donc être supprimées avec l'outil "band math" de Orfeo Tool Box. En effet, elles empêcheraient l'addition ultérieure des rasters.

La formule utilisée dans “band math” est : $im1b1 < 0 ? 0 : im1b1$

6.6) Addition des rasters (calculatrice raster.)

A cette étape, il est possible de pondérer en appliquant un coefficient à chaque raster élémentaire. Après discussion avec notre commanditaire, nous avons choisi de minimiser le poids des données SIRENE, celles-ci étant dans l'ensemble peu fiables. Par ailleurs, nous attribuons un poids relatif important à la donnée démographique. En effet le lieu de vie des habitants est un facteur essentiel dans la définition d'une polarité. D'autres part, nous donnons un poids plus important aux équipements de proximité qu'aux autres, considérant qu'ils constituent les services les plus recherchés.

Formule utilisée :

$(Population * 5) + (Sirene * 3) + (Équipements\ de\ proximité * 2) + (Équipements\ intermédiaires) + (Équipements\ supérieurs)$

6.7) Symbologie

Les valeurs du raster final sont étirées et nous appliquons un rendu en pseudo-couleur à bande unique. L'interpolation se fait en mode discret avec des seuils choisis empiriquement de manière à faire apparaître les polarités les plus modestes. Une analyse plus fine de l'histogramme des valeurs de polarité n'a pas été menée, faute de temps, mais pourrait aider au choix des seuils.

-Captures d'écran de l'onglet style pour l'indice de polarité de la Haute-Garonne (exceptées le SCoT Grande Agglomération Toulousaine) :

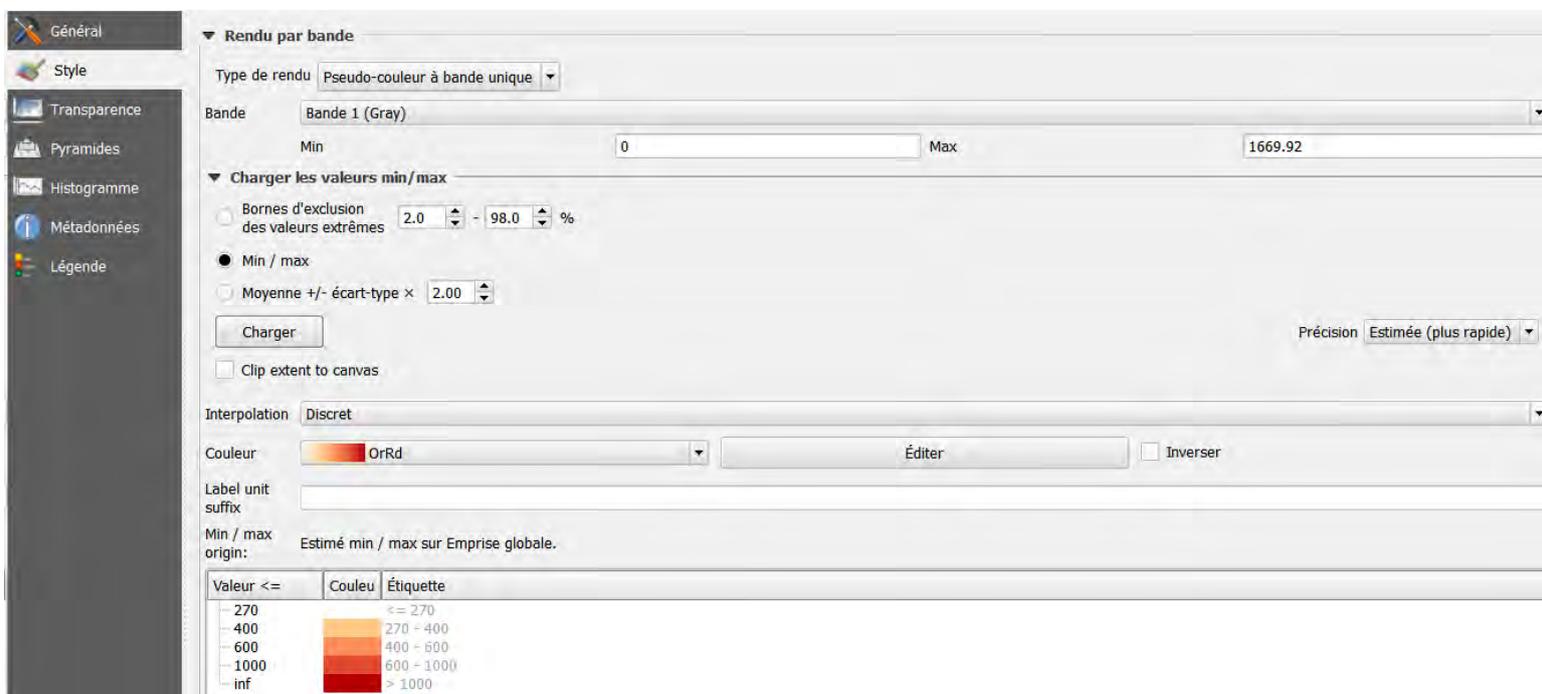


Figure 14 : Onglet style de la carte des polarités de la Haute-Garonne

- Captures d'écran de l'onglet style pour l'indice de polarité du SCoT Grande Agglomération Toulousaine :

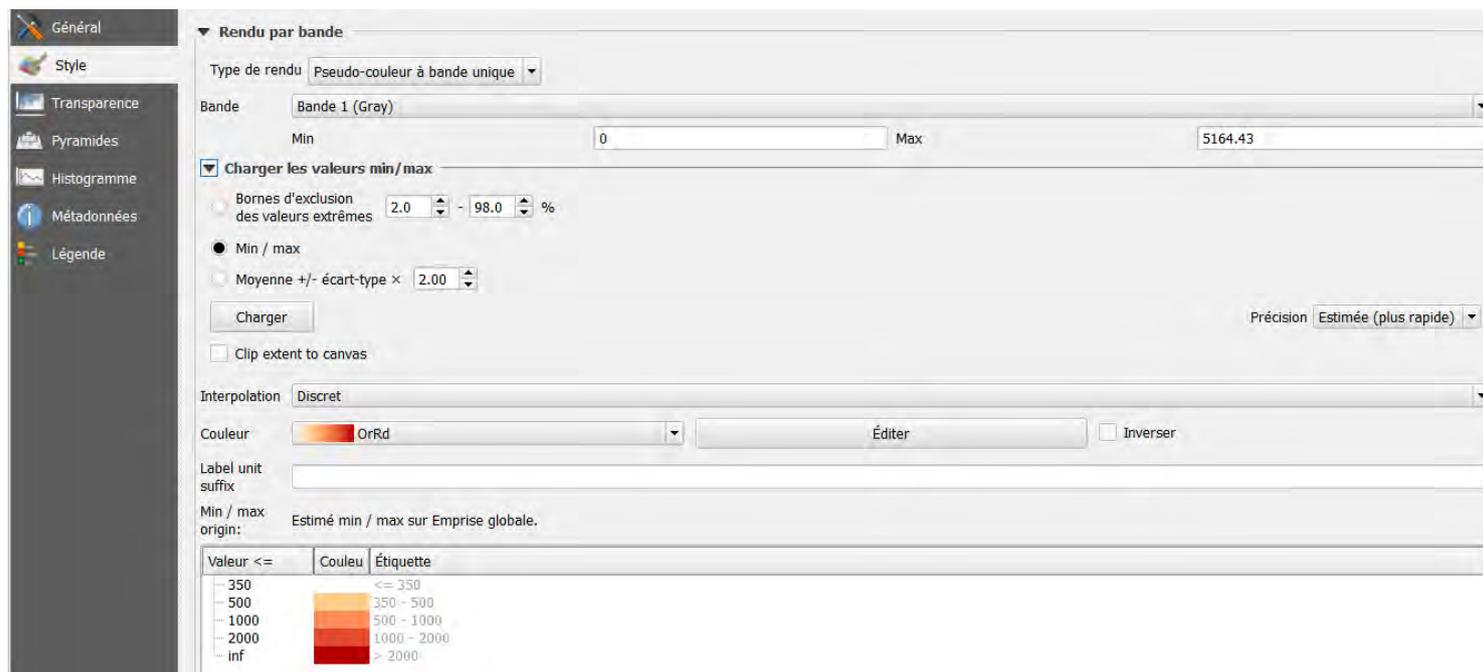


Figure 15 Onglet style de la carte des polarités du SCoT Grande agglomération Toulousaine

Cf Annexe page 69 : Diagramme de flow - Indicateurs par maille

Cf Annexe page 70 : Diagramme de flow - Indicateurs par point

Identification des polarités de la Haute-Garonne en 2015
 Hors SCoT Grande Agglomération Toulousaine - Echelle infra-communale
 Agrégation des indicateurs d'activité, d'équipements et de population

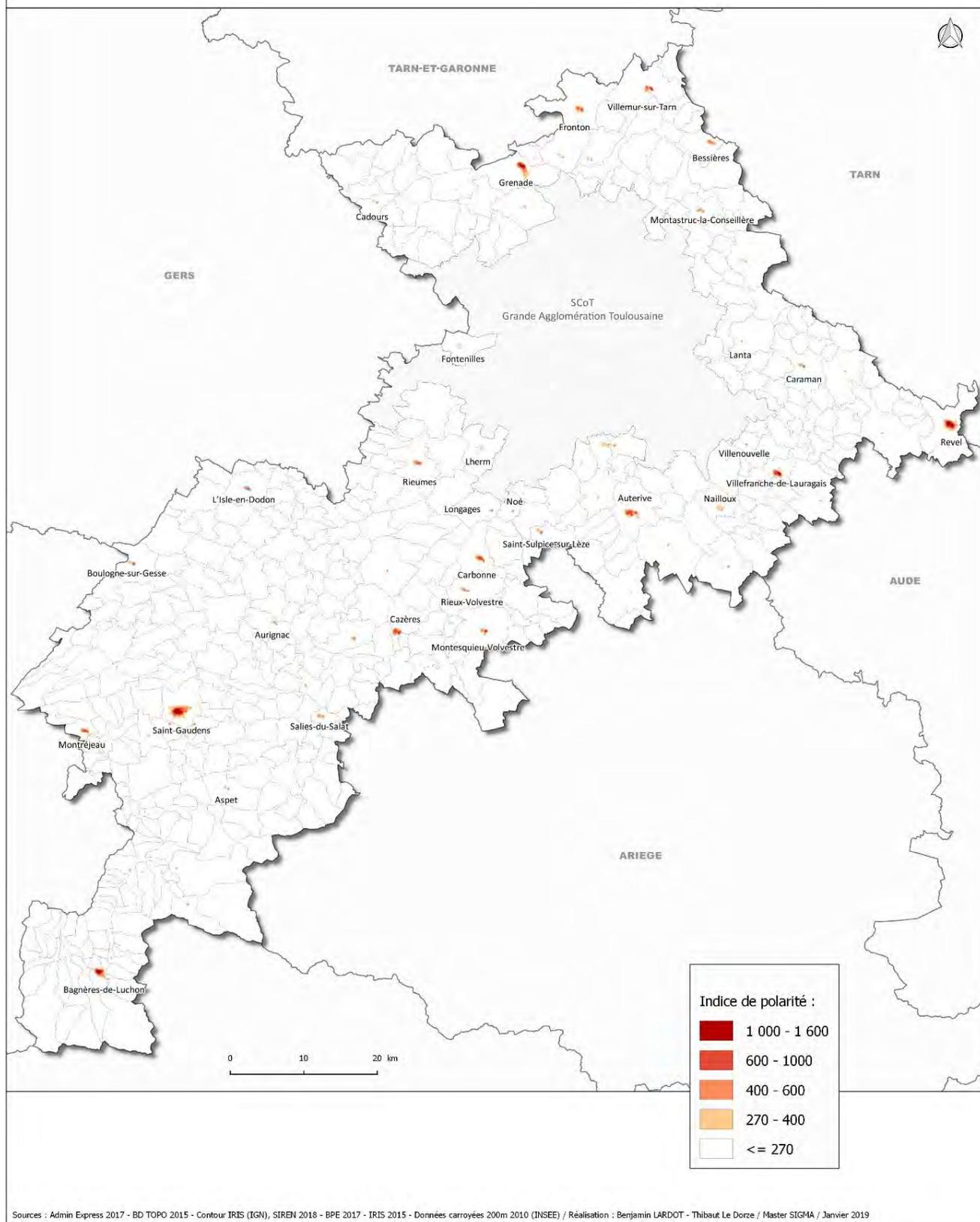


Figure 16 : Carte finale Haute-Garonne

Identification des polarités du SCoT Grande Agglomération Toulousaine en 2015 - Echelle infra-communale
Agrégation des indicateurs d'activité, d'équipements et de population

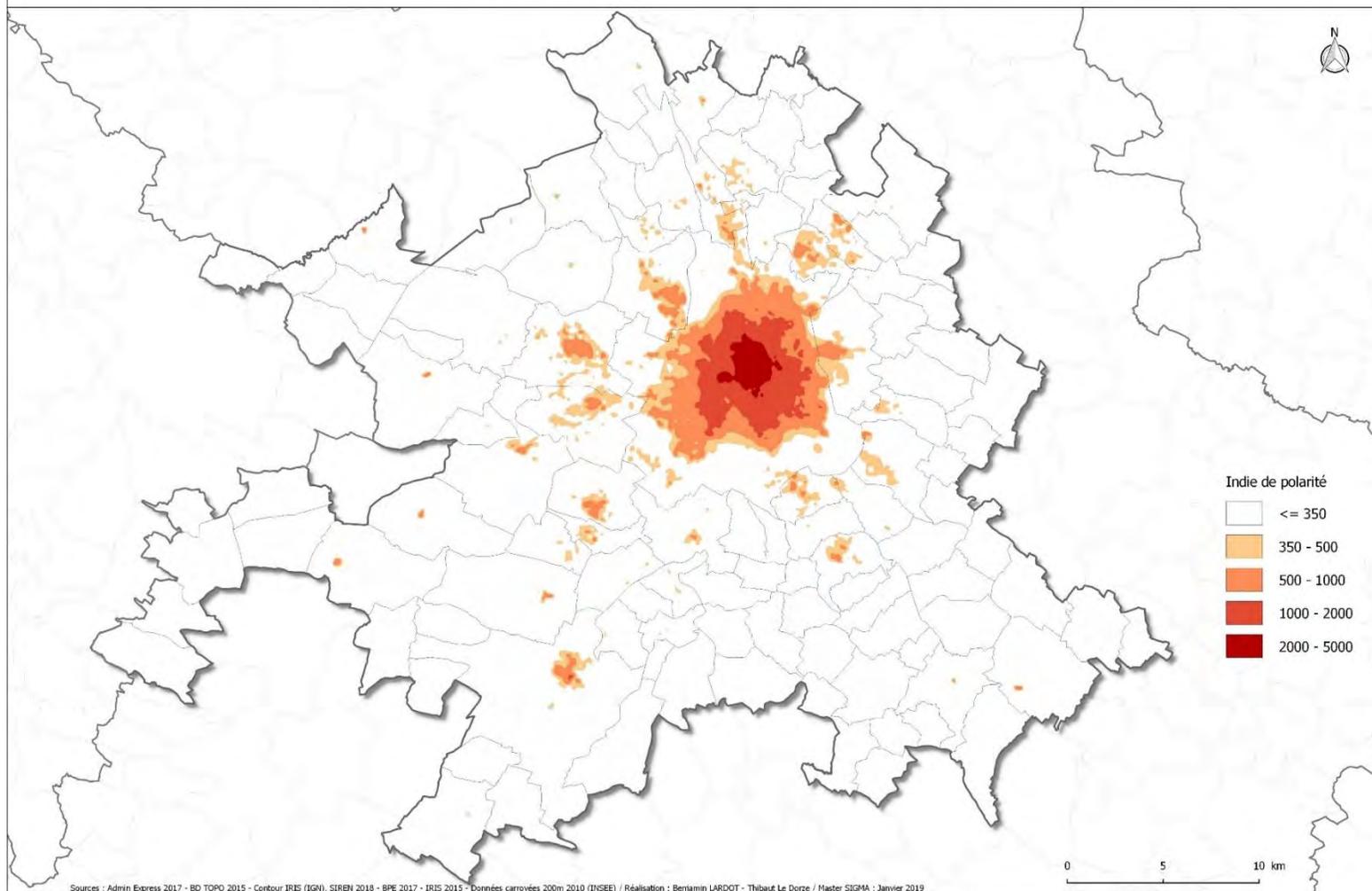


Figure 17 : Carte finale SCoT Grande Agglomération Toulousaine

NB : Le projet QGIS livré avec ce rapport, comporte les cartes finales avec les paramètres décrits ci-dessus, ainsi que d'autres exemples de paramétrages décrits ci-dessous :

-Haute-Garonne excepté le SCoT Grande Agglomération Toulousaine :

	Somme A	Somme C	Somme D	Pondération
Population	Rayon 200 m	Rayon 500 m	Rayon 200 m	5
Equipements de proximité	Rayon 200 m	Rayon 500 m	Rayon 500 m	2
Equipements intermédiaires	Rayon 600 m	Rayon 1000 m	Rayon 1000 m	1
Equipements supérieurs	Rayon 1000 m	Rayon 5000 m	Rayon 5000 m	1
Etablissements	Rayon 200 m	Rayon 300 m	Rayon 500 m	3

-SCoT Grande Agglomération Toulousaine :

	Somme 1	Somme 2	Somme 3	Pondération
Population	Rayon 100 m	Rayon 250 m	Rayon 200 m	5
Equipements de proximité	Rayon 250 m	Rayon 400 m	Rayon 500 m	2
Equipements intermédiaires	Rayon 400 m	Rayon 1000 m	Rayon 1000 m	1
Equipements supérieurs	Rayon 1000 m	Rayon 5000 m	Rayon 5000 m	1
Etablissements	Rayon 200 m	Rayon 400 m	Rayon 500 m	3

Figure 18 : Paramètres cartes finales : réglages figurant dans le projet QGIS

7) Retour critique sur le projet

Apports	Limites
Collaboration avec différents partenaires pour les arbitrages méthodologique (DDT) et l'aide technique (Université)	Manque de cadrage sur la définition du sujet d'étude
Gestion de l'organisation personnelle pour répondre à la commande	Manque de liberté sur certains choix méthodologiques vis-à-vis des enseignants
Etude faisant intervenir des thématiques variées avec une diversité de données	Non accessibilité des données, notamment pour l'étude synchronique en 2015 (BPE et SIRENE)
Meilleure connaissance des données existantes et des fournisseurs mobilisés	Manque de fiabilité des données, principalement pour la BPE et SIRENE
Capacité d'adaptation pour gérer de la donnée volumineuse	Temps de calcul importants pour effectuer les traitements à l'échelle du département
Utilisation de nouveaux outils géomatiques ou amélioration dans leur pratique : cartes de chaleurs, calculatrice raster, requêtes SQL...	Délai imparti limité pour la réalisation du projet qui n'a pas permis d'approfondir l'étude

Figure 19 : Tableau des apports et des limites du projet

8) Résultats et limites

L'indice de polarité généré couvre l'ensemble du département avec une résolution de 20 mètres. Les cartes obtenues sont satisfaisantes même si elles ne permettent pas de d'identifier une véritable hiérarchie urbaine au sein du département.

Les résultats sont meilleurs quand on exclut les données du SCoT de la métropole car celles-ci représentent environ $\frac{1}{3}$ de l'ensemble des données et sont très denses. En effet, les résultats sont plus diffus car la répartition spatiale des données est particulièrement dense. En revanche sur le territoire hors SCoT de Toulouse, des polarités relativement faibles dans les environs de Saint-Gaudens sont bien détectées.

Bien que les limites des données sources composant chaque indicateur aient été précisées en amont, il subsiste d'autres limites concernant l'indice de polarité devant être prises en compte pour son interprétation.

L'indice de polarité étant représenté par points, sa fiabilité est intrinsèquement dépendante de la qualité de géolocalisation des données. Pour limiter ce biais, la BPE et la base SIRENE utilisées respectivement pour les indicateurs d'équipements et d'activité ont été filtrées en ne retenant que les points ayant le meilleur géocodage, avec une pondération additionnelle pour la base SIRENE selon la valeur du géo-score.

Malgré ces précautions, on peut constater par exemple que des polarités mineures détectées dans le secteur du Comminges correspondent en réalité à une accumulation d'établissements géolocalisés au centroïde de leur commune d'appartenance (et non selon leur localisation réelle).

D'autre part, les résultats obtenus doivent être nuancés au regard d'autres polarités éventuelles situées dans les départements limitrophes pouvant influencer sur les polarités présentes en Haute-Garonne compte-tenu de l'effet de frontière. En effet, si l'importance d'une polarité dépend de sa densité, elle dépend également de l'éloignement d'autres polarités. Ainsi, la proximité de polarités fortes situées en dehors du département peut freiner le développement des polarités visibles en Haute-Garonne.

La méthode et les valeurs de pondération attribuées à l'indice de polarité peuvent également être nuancées car elles sont intrinsèquement subjectives et influent grandement sur son calcul. Il est possible de les modifier pour adapter l'importance à un indicateur selon la finalité du rendu.

9) Perspectives

Le commanditaire souhaitant un rendu sous forme de carte de chaleur, l'indice de polarité n'a pas été injecté dans le carroyage INSEE mais cette opération est réalisable. Cela serait d'autant plus intéressant que l'INSEE devrait à mettre à disposition une nouvelle version de la donnée carroyée courant mai 2019.

Il est ainsi possible de s'affranchir d'estimation de la population pour l'indicateur démographique et de fiabiliser l'indice de polarité, bien que l'on perde en précision par rapport à une représentation au point. Par ailleurs, la BPE est mise à jour annuellement et la base SIRENE en continu ce qui permettrait d'utiliser des données d'une même année pour chaque indicateur thématique et d'éviter le biais méthodologique présent au calcul de cet indice de polarité.

Néanmoins des limites subsistent sur cette donnée dans l'absence de communication officielle de l'INSEE quant au contenu des données attributaires, à la complétude du carroyage, à sa résolution et à sa fréquence de mise à jour.

Une alternative est de ne pas utiliser la donnée carroyée en reproduisant l'estimation de population selon la méthode décrite dans le mémoire technique. Les principaux avantages sont de pouvoir effectuer une mise à jour annuelle de l'indice de polarité en utilisant le redressement annuel du recensement de la population et de disposer d'une résolution plus précise pour la représentation.

La reproduction de cette étude avec des données actualisées en 2019 sera donc possible d'ici à la réédition de la BPE en courant juillet 2020. Une première comparaison pourra être effectuée pour étudier l'évolution des polarités en Haute-Garonne, mais elle sera à nuancer au regard des ajustements introduits dans la méthodologie et l'intégration des données.

Afin de fiabiliser davantage l'indice de polarité, il est envisageable d'intégrer des données issues de fournisseurs locaux ou d'enquêtes de terrain à la demande d'un établissement public comme la DDT 31, mais dont le coût dissuasif pourrait restreindre le secteur d'application. Par ailleurs, ce changement de méthodologie risquerait d'impacter le critère de reproductibilité sur d'autres territoires souhaités dans cette étude et de rendre des comparaisons impossibles.

10) Conclusion

Ces résultats restent donc perfectibles et doivent être interprétés en prenant en compte les biais intervenant dans les différents stades de création de l'indice de polarité. Cependant avec les données disponibles, cette étude a permis de fixer un état des lieux des polarités présentes en Haute-Garonne sur l'année de référence 2015.

Dans une application plus poussée et avec des données plus fiables, l'indice de polarité permettrait de mieux comprendre l'organisation et les dynamiques territoriales et de profiler l'armature urbaine présente en Haute-Garonne. Les pouvoirs publics peuvent s'emparer de cet outil comme une aide à la décision opérationnelle afin d'orienter leurs politiques et de contribuer à un aménagement du territoire plus équilibré.

11) Annexes

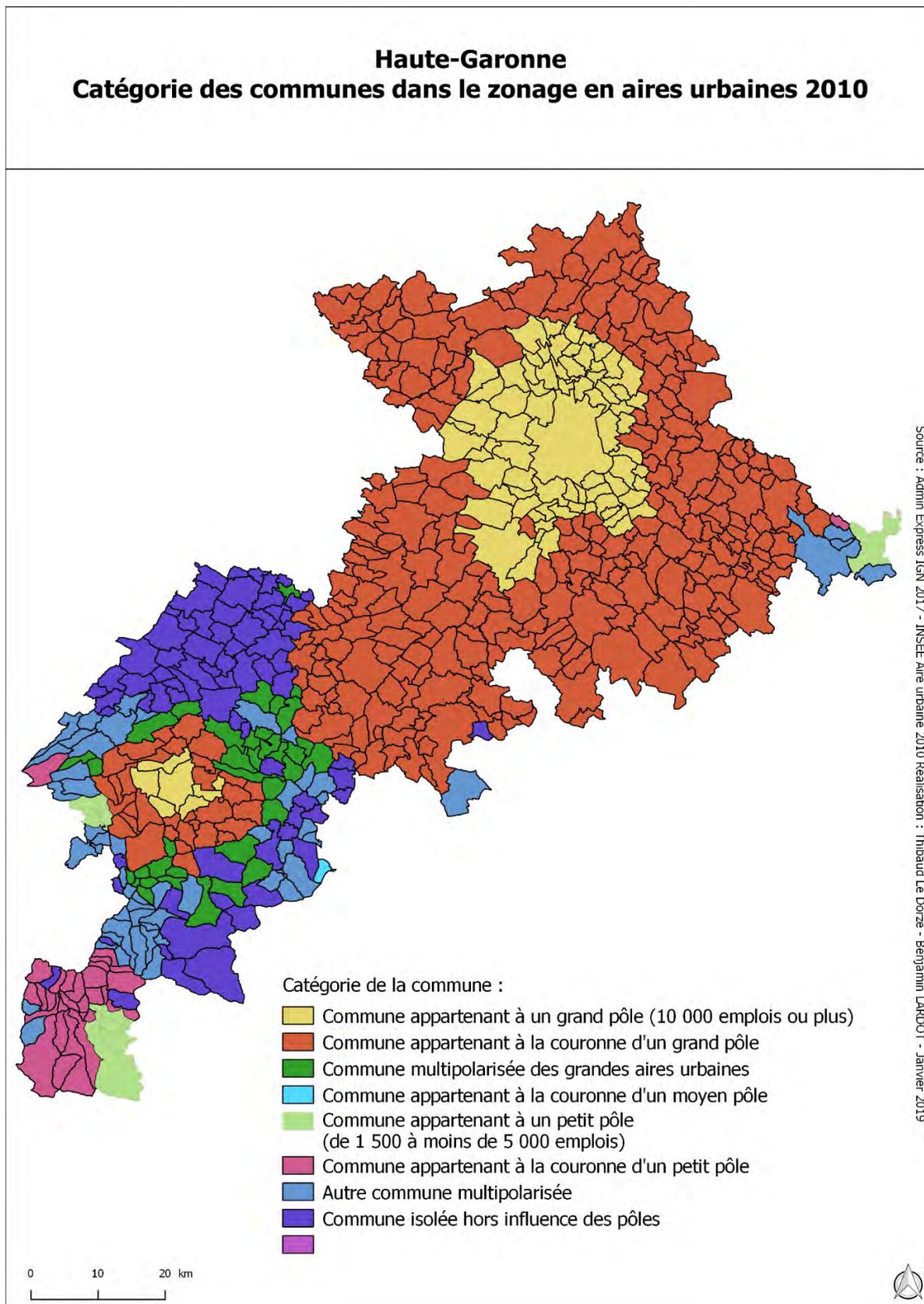


Figure 20 : Carte de zonage en aires urbaines

Haute-Garonne Evolution de la population communale entre 1975 et 2015

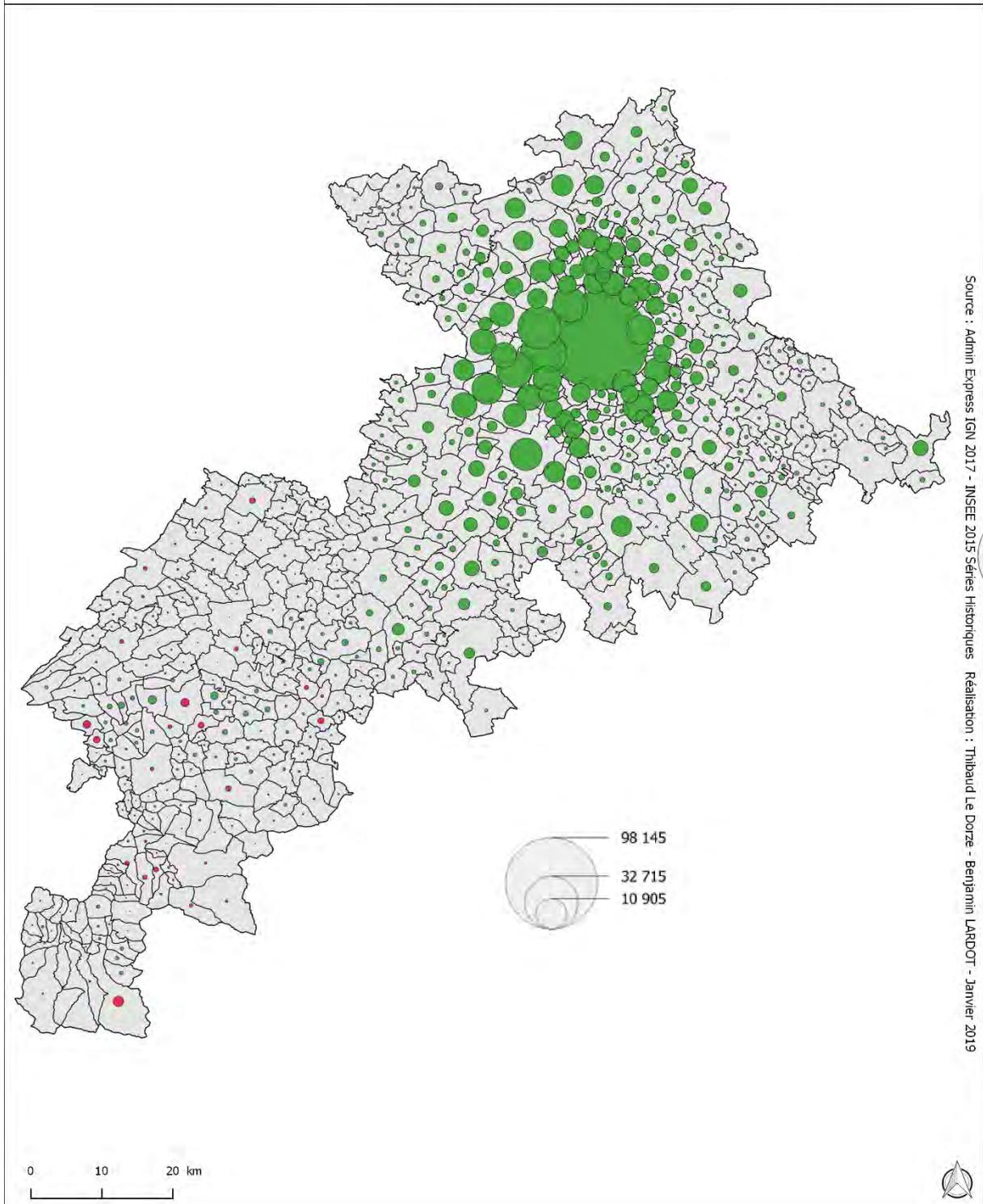


Figure 21 : Carte de l'évolution de la population communale 1975 – 2015

Haute-Garonne - 2015

Nombre d'établissements du secteur privé par commune

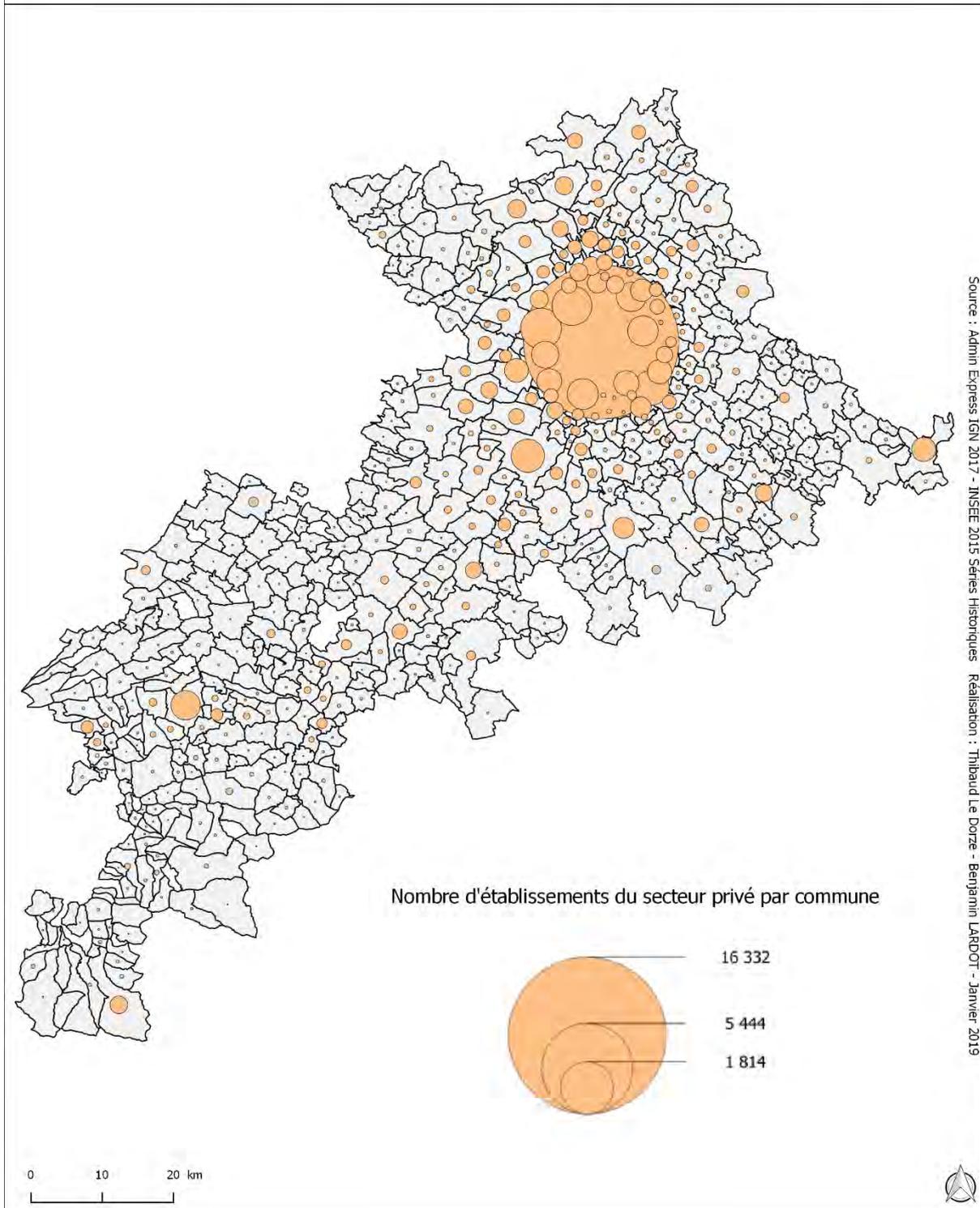
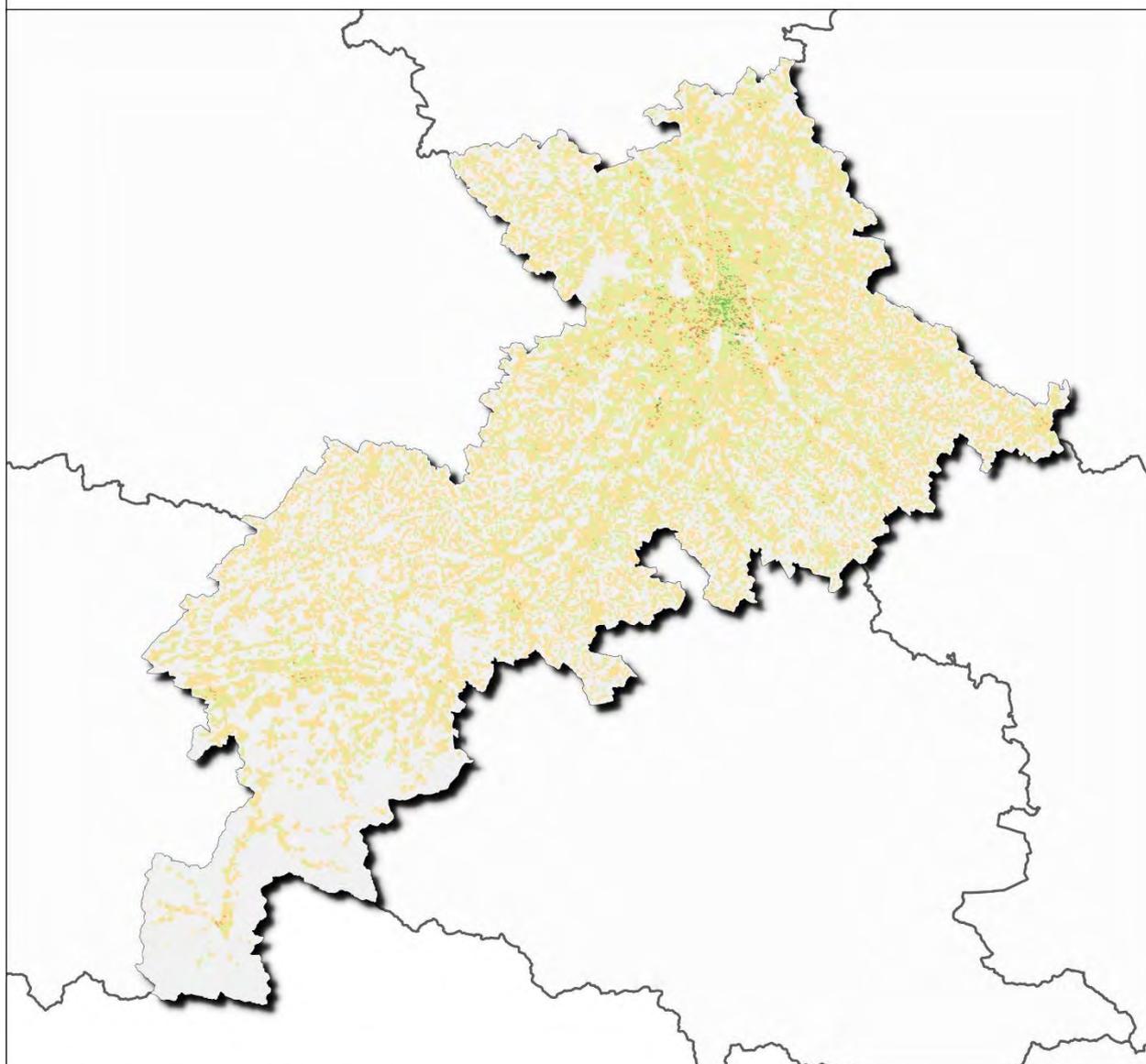
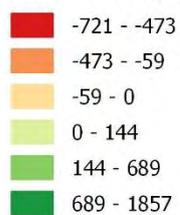


Figure 22 : Carte du nombre d'établissement du secteur privé par commune

Evolution de la population entre 2010 et 2015
dans le carroyage INSEE
Département de Haute-Garonne



Evolution du nombre d'habitants
2010 - 2015 par maille :



0 10 20 km

Sources : Admin Express IGN 2017 - BD TOPO IGN 2015 - IRIS 2015 INSEE - Carroyage INSEE 200 m 2010 Réalisation : Benjamin LARDOT - Thibaut Le Dorze

Janvier 2019

Figure 23 : Carte d'évolution de la population entre 2010 et 2015 dans le carroyage INSEE

Diagramme de flow - Etablissements

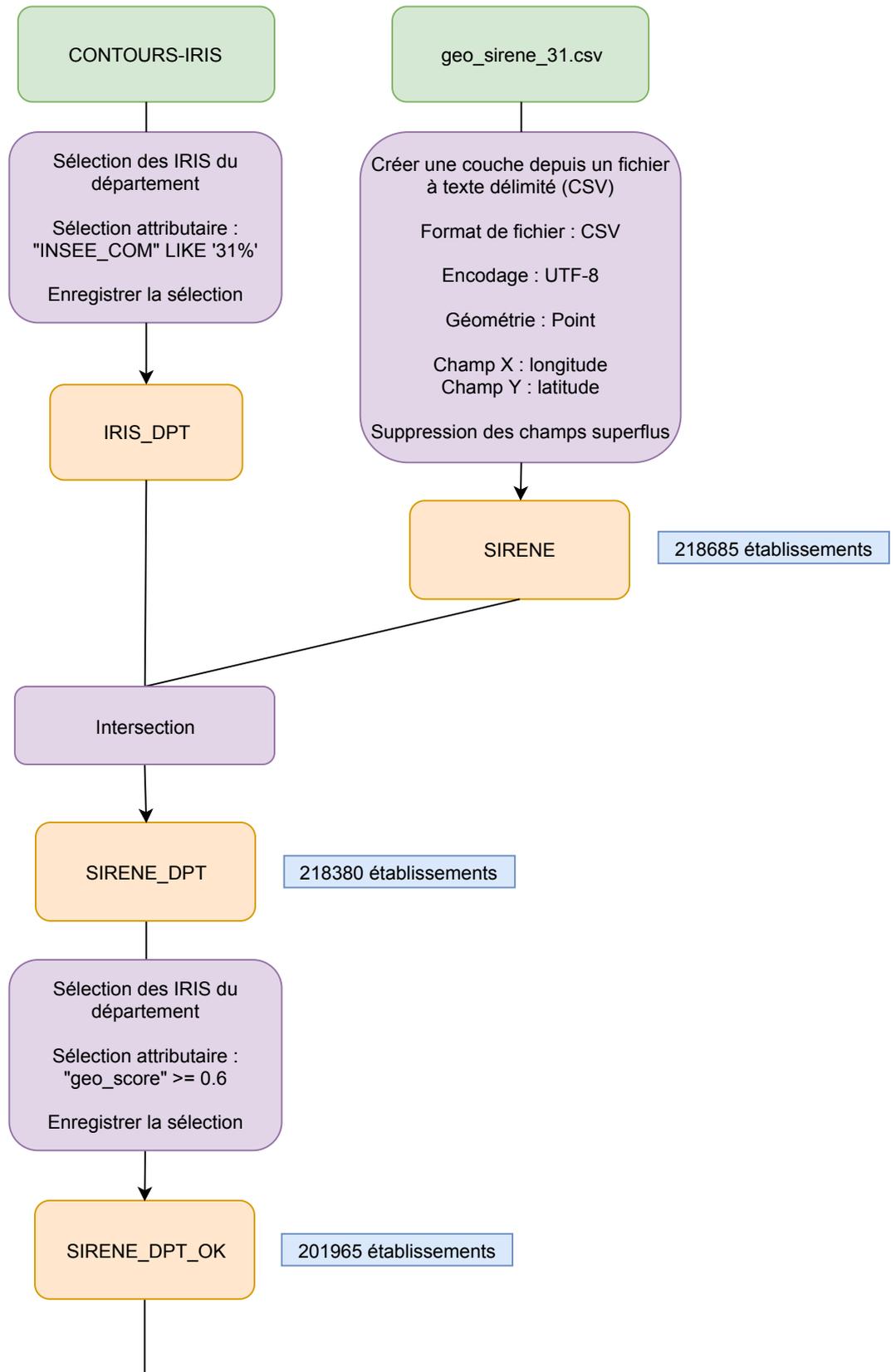
Légende

Donnée source

Couche intermédiaire

Traitement

Indicateur



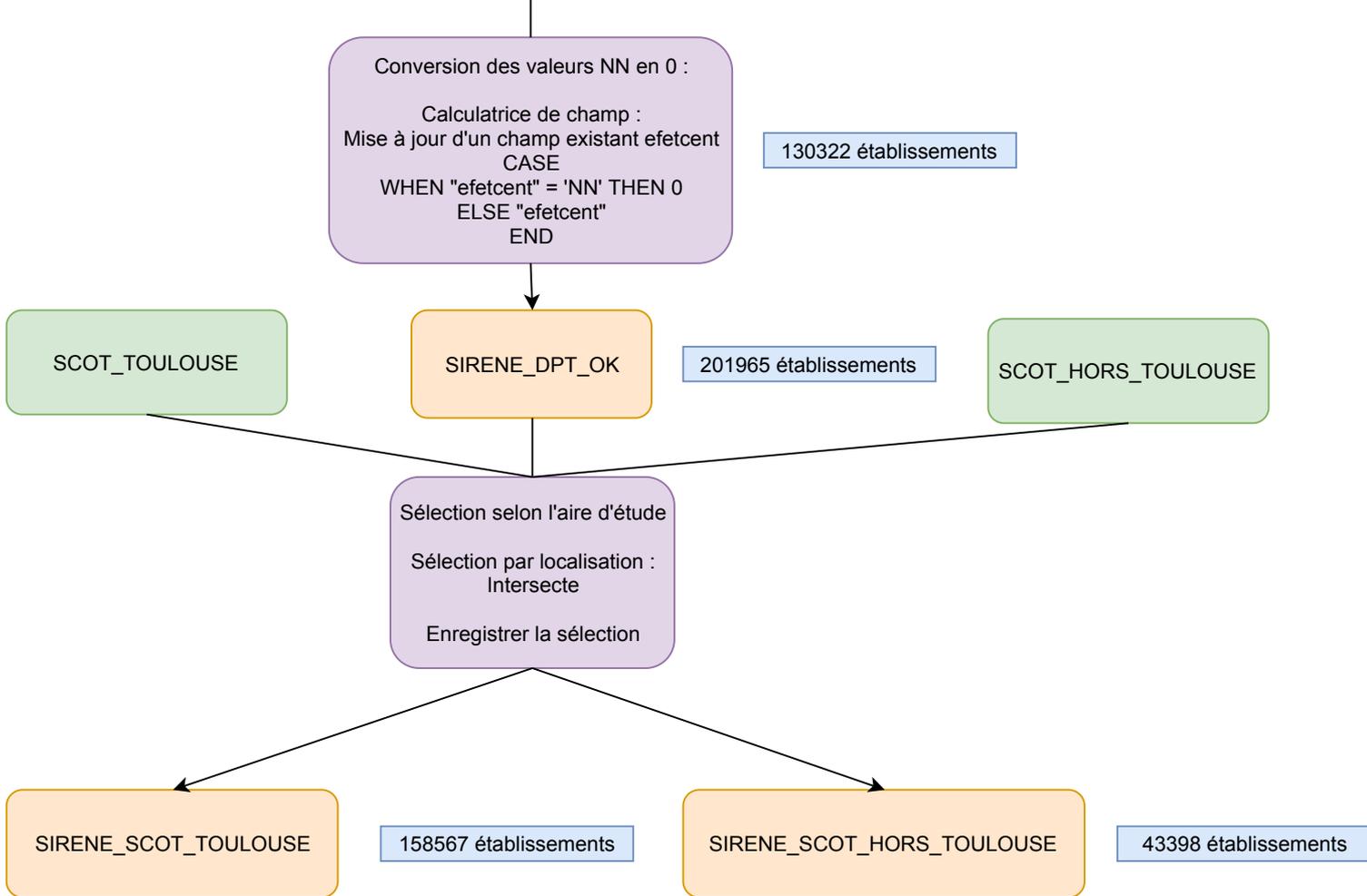


Diagramme de flow - Population par bâtiment

Légende

Méthode LCSQA

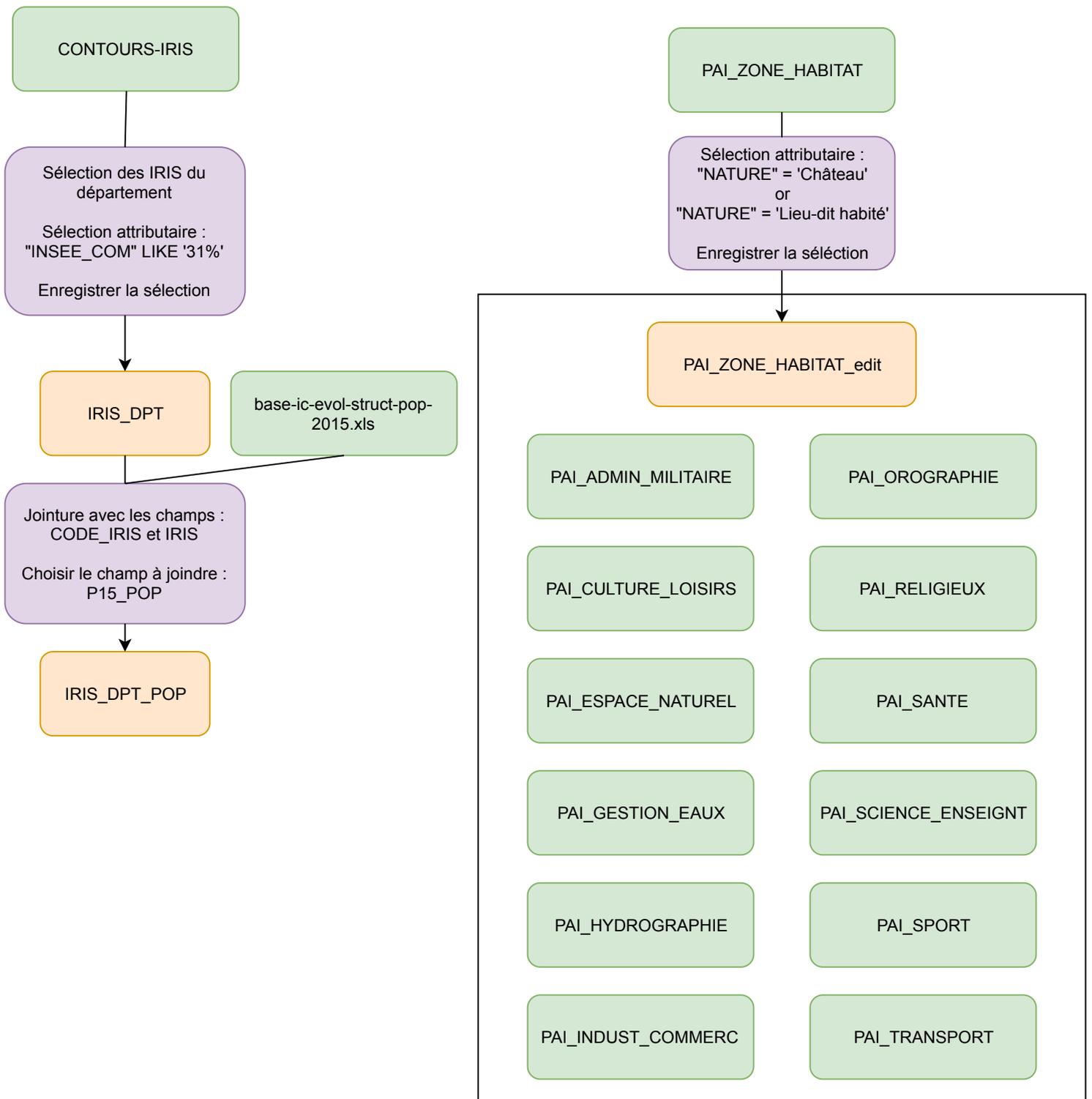
Donnée source

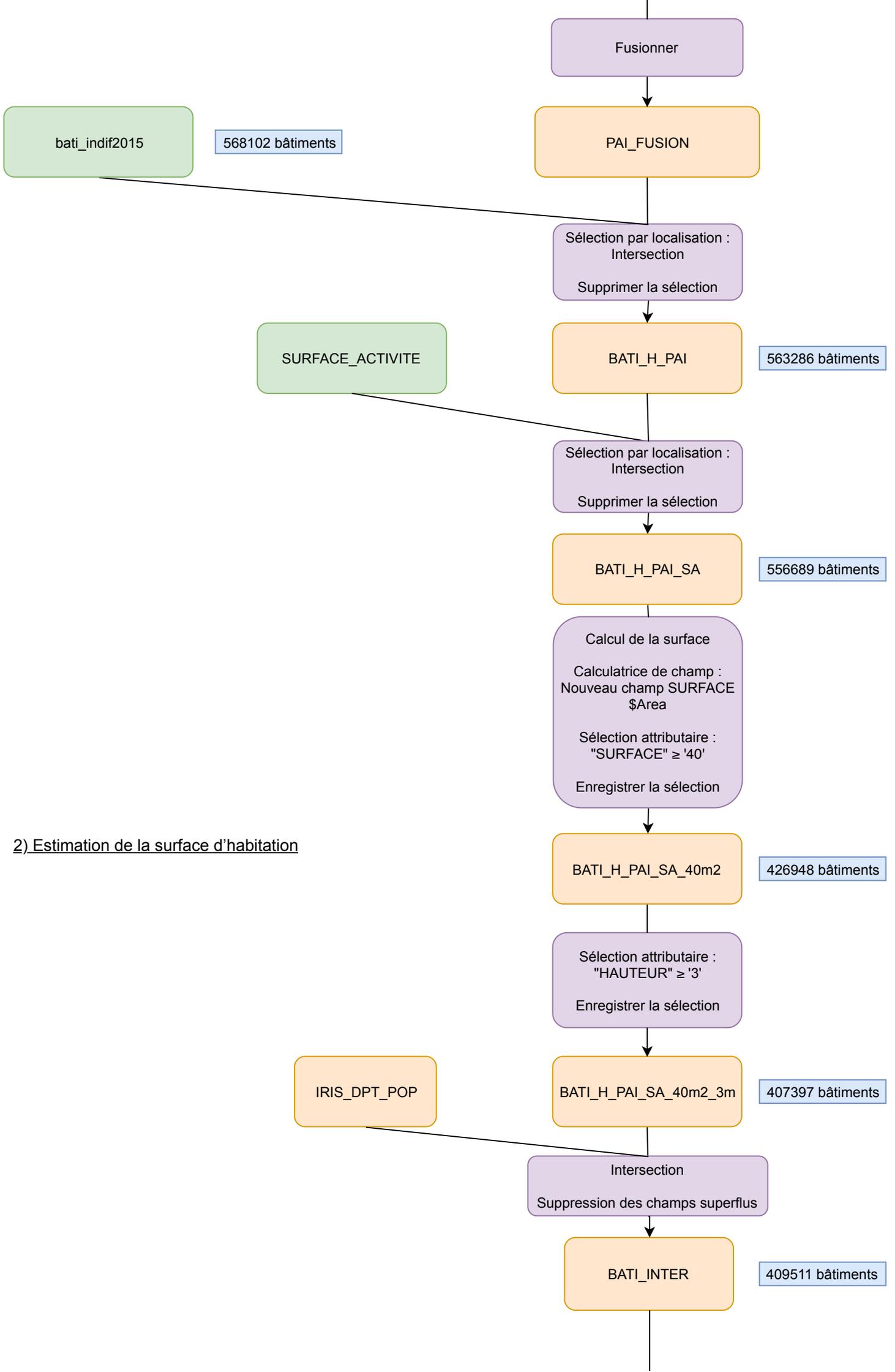
Couche intermédiaire

Traitement

Indicateur

1) Sélection du bâti résidentiel





```
Calcul du nombre d'étages

Calculatrice de champ :
Nouveau champ NB_ETAGES
CASE
WHEN "HAUTEUR" < 6 THEN 1
WHEN ("HAUTEUR" >= 6 and "HAUTEUR"<9) THEN 2
WHEN ("HAUTEUR" >= 9 and "HAUTEUR"<12) THEN 3
WHEN ("HAUTEUR" >= 12 and "HAUTEUR"<15) THEN 4
WHEN ("HAUTEUR" >= 15 and "HAUTEUR"<18) THEN 5
WHEN ("HAUTEUR" >= 18 and "HAUTEUR"<21) THEN 6
WHEN ("HAUTEUR" >= 21 and "HAUTEUR"<24) THEN 7
WHEN ("HAUTEUR" >= 24 and "HAUTEUR"<27) THEN 8
WHEN ("HAUTEUR" >= 27 and "HAUTEUR"<30) THEN 9
WHEN ("HAUTEUR" >= 30 and "HAUTEUR"<33) THEN 10
WHEN ("HAUTEUR" >= 33 and "HAUTEUR"<36) THEN 11
WHEN ("HAUTEUR" >= 36 and "HAUTEUR"<39) THEN 12
WHEN ("HAUTEUR" >= 39 and "HAUTEUR"<42) THEN 13
WHEN ("HAUTEUR" >= 42 and "HAUTEUR"<45) THEN 14
WHEN ("HAUTEUR" >= 45 and "HAUTEUR"<48) THEN 15
WHEN ("HAUTEUR" >= 48 and "HAUTEUR"<51) THEN 16
WHEN ("HAUTEUR" >= 51 and "HAUTEUR"<54) THEN 17
WHEN ("HAUTEUR" >= 54 and "HAUTEUR"<57) THEN 18
WHEN ("HAUTEUR" >= 57 and "HAUTEUR"<60) THEN 19
END
```

```
Calcul de la surface habitable

Calculatrice de champ :
Nouveau champ SURF_INTER
$Area

Calculatrice de champ :
Nouveau champ SURF_HAB
SURF_INTER*NB_ETAGES
```

BATI_SURF_HAB

```
Sélection de la surface habitable des bâtiments contenus dans chaque IRIS

Gestionnaire de base de données :
select BATI_SURF_HAB.*, sum(BATI_SURF_HAB.SURF_HAB) as SURF_HAB_I, BATI_SURF_HAB.geometry
from BATI_SURF_HAB
group by BATI_SURF_HAB.CODE_IRIS
```

SURF_HAB_I

```
Jointure avec le champ :
CODE_IRIS

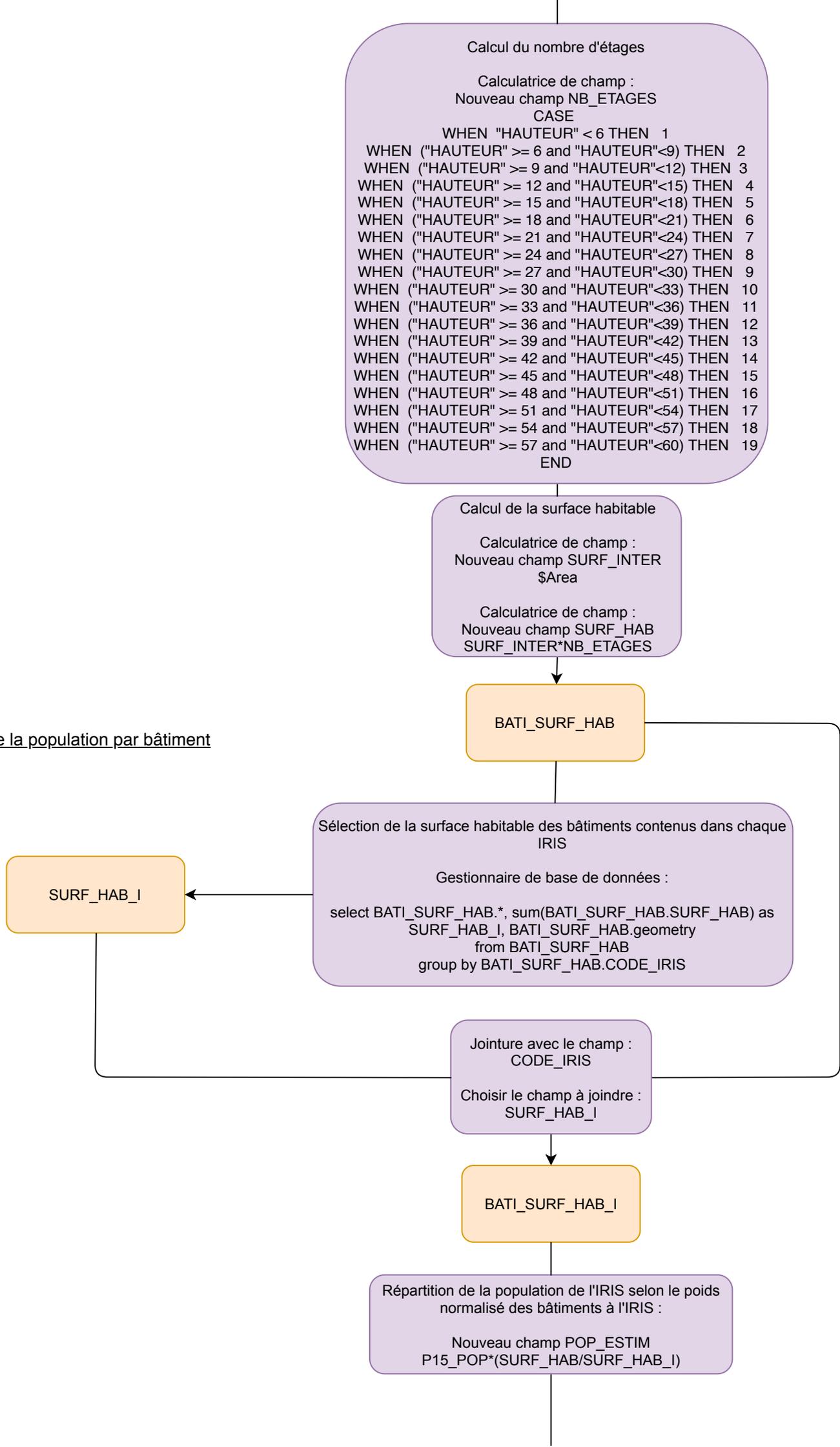
Choisir le champ à joindre :
SURF_HAB_I
```

BATI_SURF_HAB_I

```
Répartition de la population de l'IRIS selon le poids
normalisé des bâtiments à l'IRIS :

Nouveau champ POP_ESTIM
P15_POP*(SURF_HAB/SURF_HAB_I)
```

3) Estimation de la population par bâtiment



4) Sélection de l'aire d'étude

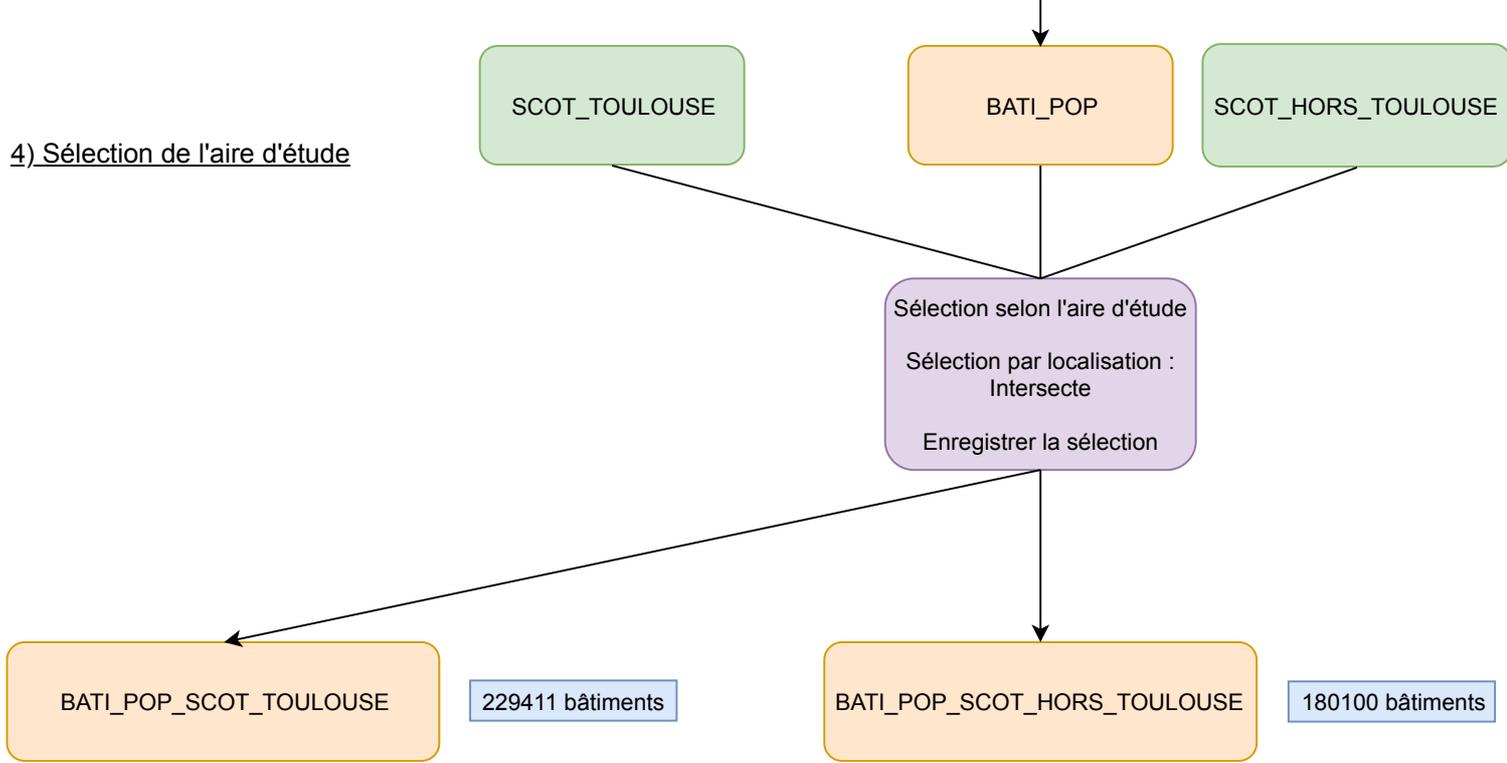
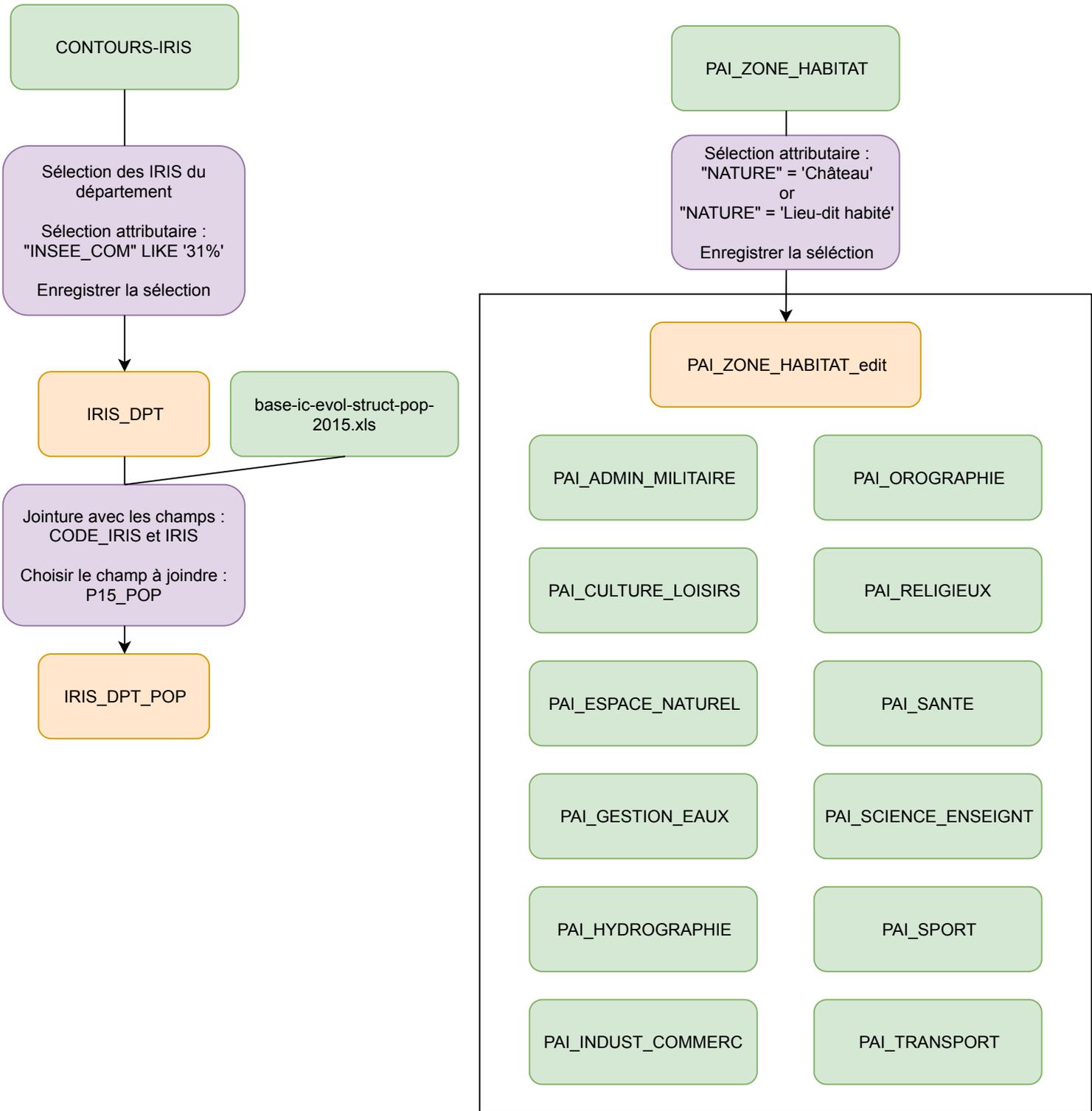


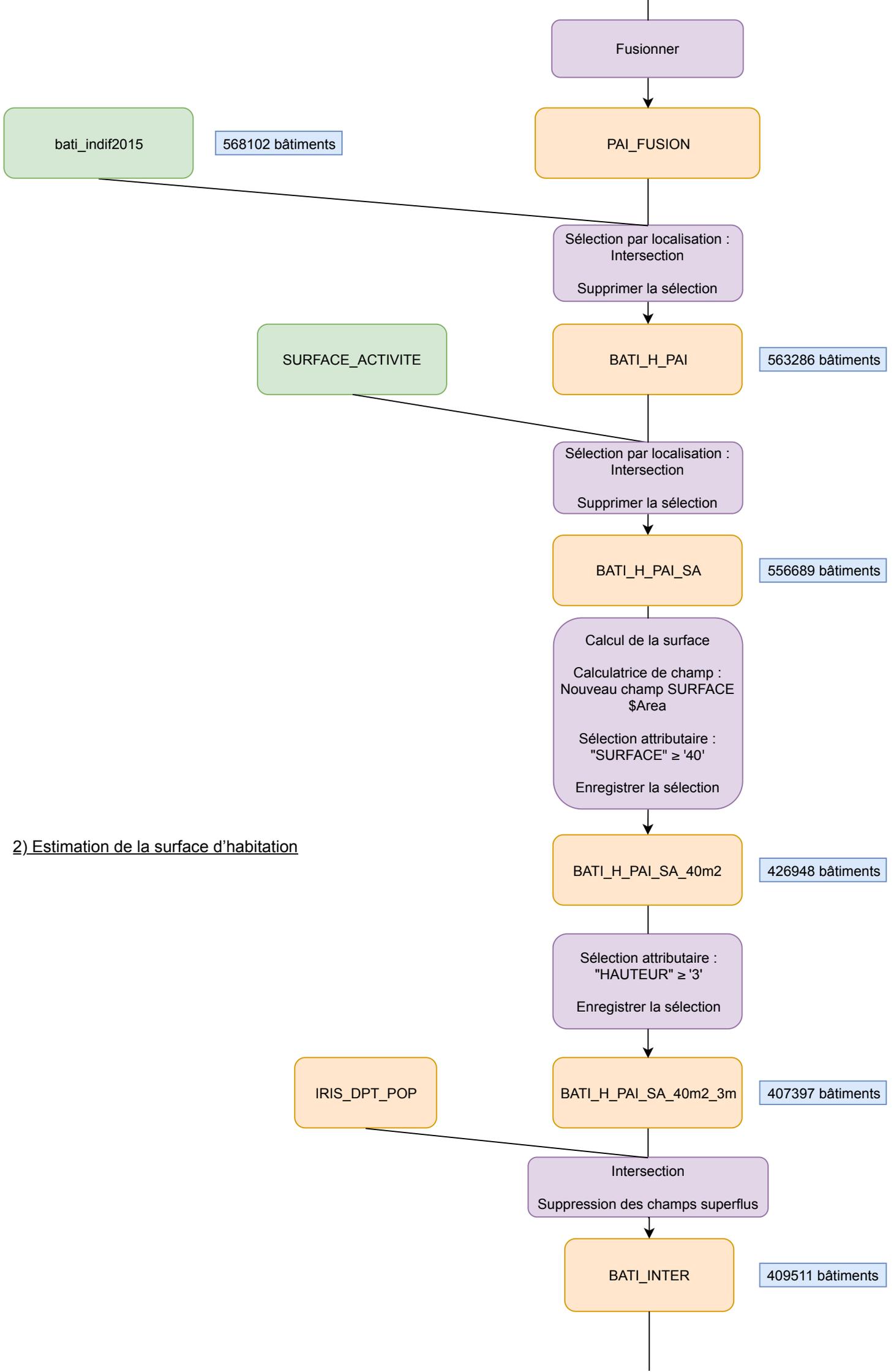
Diagramme de flow - Population par maille

Légende
Méthode LCSQA

- Donnée source
- Couche intermédiaire
- Traitement
- Indicateur

1) Sélection du bâti résidentiel





```
Calcul du nombre d'étages

Calculatrice de champ :
Nouveau champ NB_ETAGES
CASE
  WHEN "HAUTEUR" < 6 THEN 1
  WHEN ("HAUTEUR" >= 6 and "HAUTEUR"<9) THEN 2
  WHEN ("HAUTEUR" >= 9 and "HAUTEUR"<12) THEN 3
  WHEN ("HAUTEUR" >= 12 and "HAUTEUR"<15) THEN 4
  WHEN ("HAUTEUR" >= 15 and "HAUTEUR"<18) THEN 5
  WHEN ("HAUTEUR" >= 18 and "HAUTEUR"<21) THEN 6
  WHEN ("HAUTEUR" >= 21 and "HAUTEUR"<24) THEN 7
  WHEN ("HAUTEUR" >= 24 and "HAUTEUR"<27) THEN 8
  WHEN ("HAUTEUR" >= 27 and "HAUTEUR"<30) THEN 9
  WHEN ("HAUTEUR" >= 30 and "HAUTEUR"<33) THEN 10
  WHEN ("HAUTEUR" >= 33 and "HAUTEUR"<36) THEN 11
  WHEN ("HAUTEUR" >= 36 and "HAUTEUR"<39) THEN 12
  WHEN ("HAUTEUR" >= 39 and "HAUTEUR"<42) THEN 13
  WHEN ("HAUTEUR" >= 42 and "HAUTEUR"<45) THEN 14
  WHEN ("HAUTEUR" >= 45 and "HAUTEUR"<48) THEN 15
  WHEN ("HAUTEUR" >= 48 and "HAUTEUR"<51) THEN 16
  WHEN ("HAUTEUR" >= 51 and "HAUTEUR"<54) THEN 17
  WHEN ("HAUTEUR" >= 54 and "HAUTEUR"<57) THEN 18
  WHEN ("HAUTEUR" >= 57 and "HAUTEUR"<60) THEN 19
END
```

```
Calcul de la surface habitable

Calculatrice de champ :
Nouveau champ SURF_INTER
$Area

Calculatrice de champ :
Nouveau champ SURF_HAB
SURF_INTER*NB_ETAGES
```

BATI_SURF_HAB

3) Estimation de la population par bâtiment

```
Sélection de la surface habitable des bâtiments contenus dans chaque IRIS

Gestionnaire de base de données :
select BATI_SURF_HAB.*, sum(BATI_SURF_HAB.SURF_HAB) as
SURF_HAB_I, BATI_SURF_HAB.geometry
from BATI_SURF_HAB
group by BATI_SURF_HAB.CODE_IRIS
```

SURF_HAB_I

```
Jointure avec le champ :
CODE_IRIS

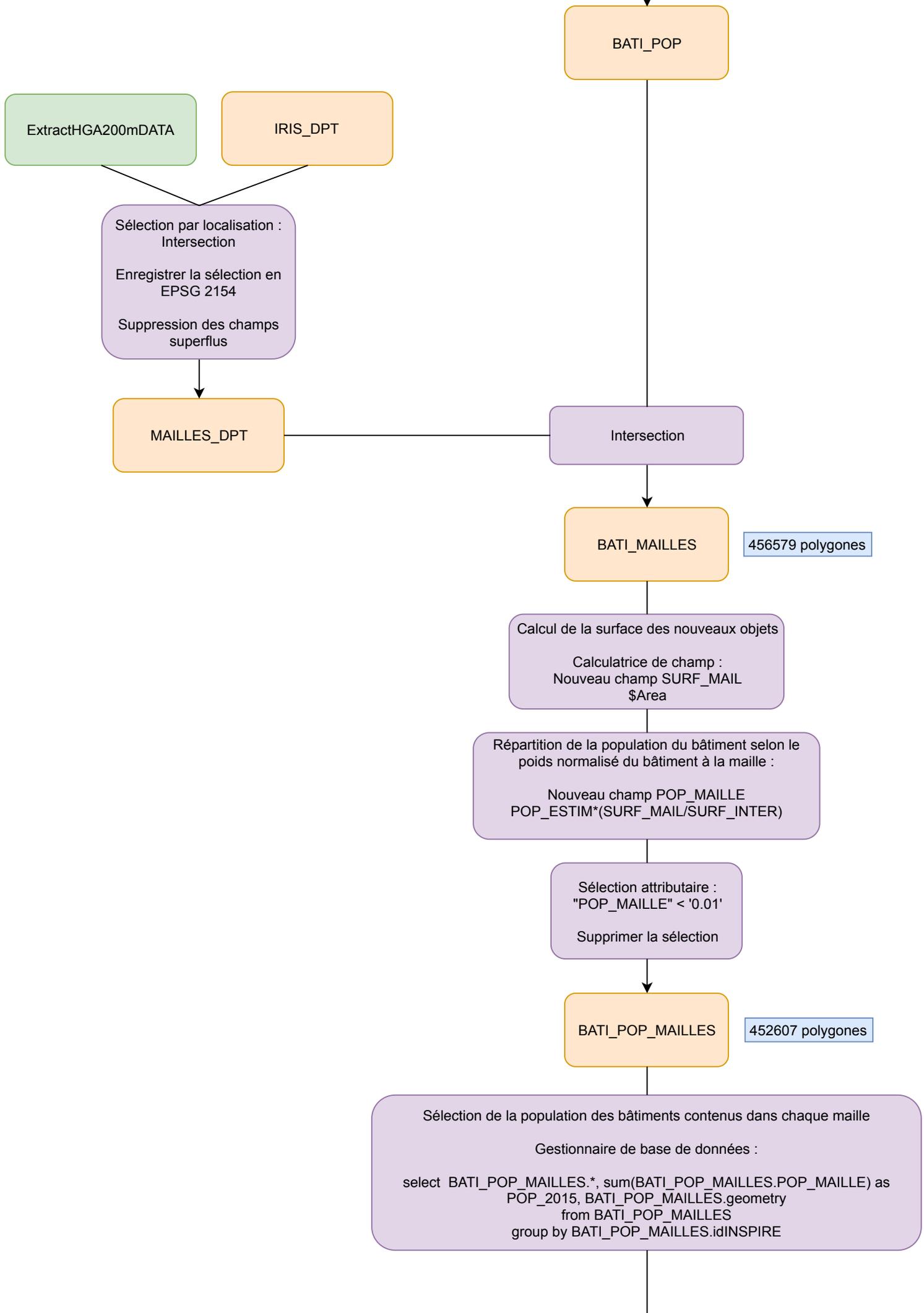
Choisir le champ à joindre :
SURF_HAB_I
```

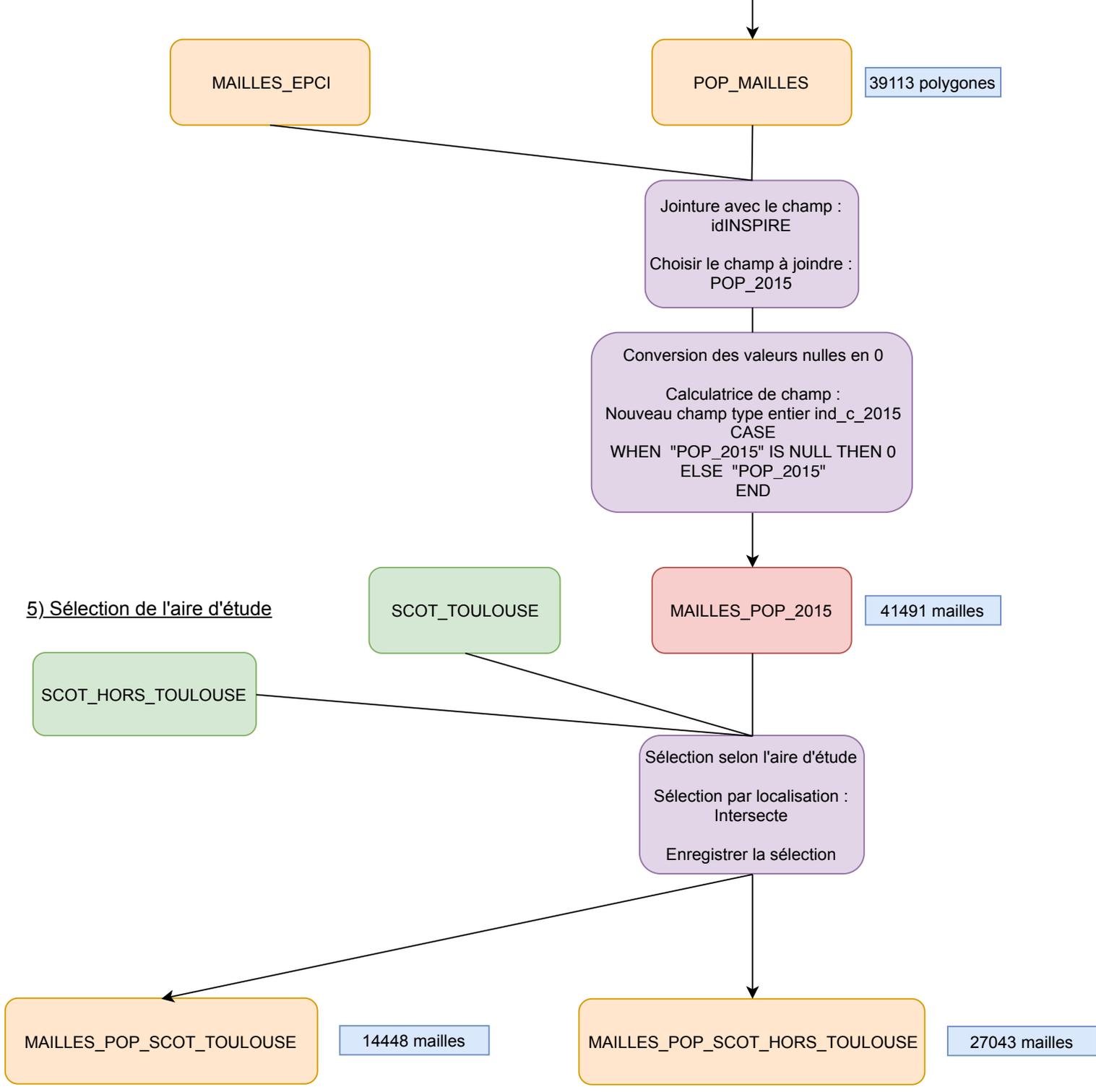
BATI_SURF_HAB_I

```
Répartition de la population de l'IRIS selon le poids
normalisé des bâtiments à l'IRIS :

Nouveau champ POP_ESTIM
P15_POP*(SURF_HAB/SURF_HAB_I)
```

4) Distribution de la population des bâtiments par maille et sélection de l'aire d'étude :





Légende :

Données source

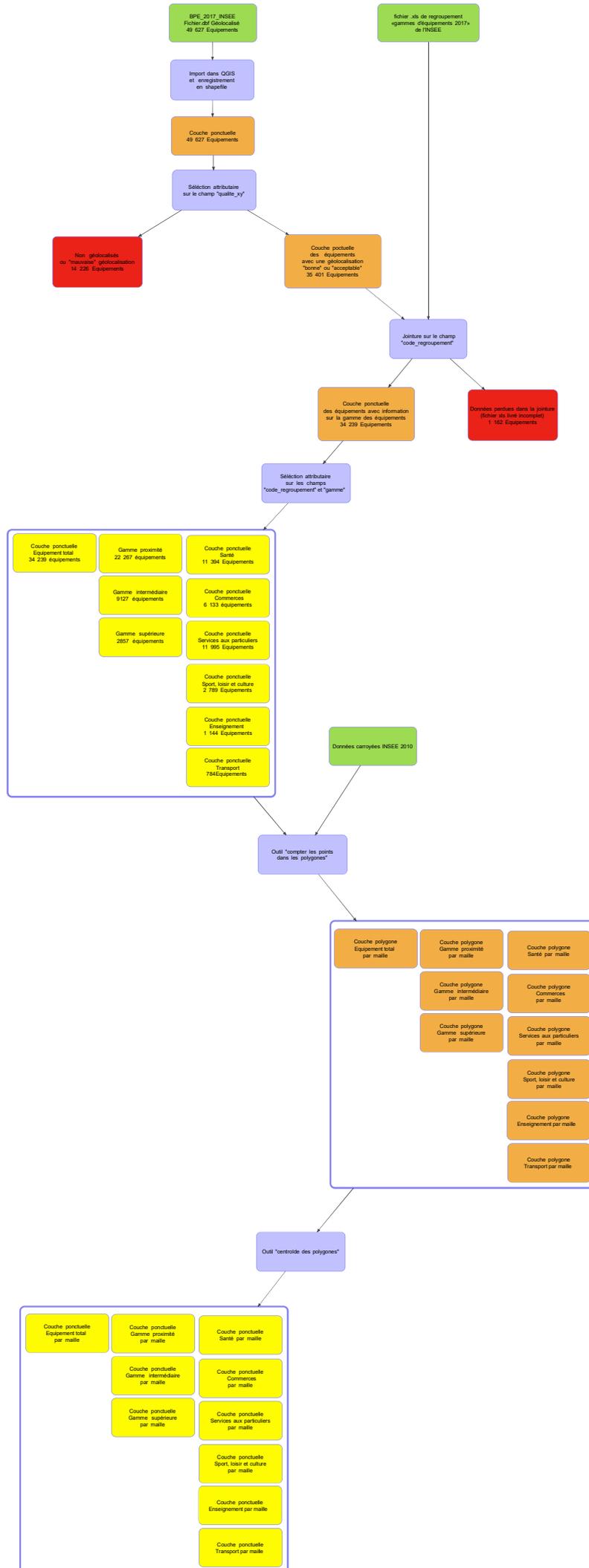
Traitement

Couche intermédiaire

Données finales pour l'outil "carte de chaleur"

Données écartées

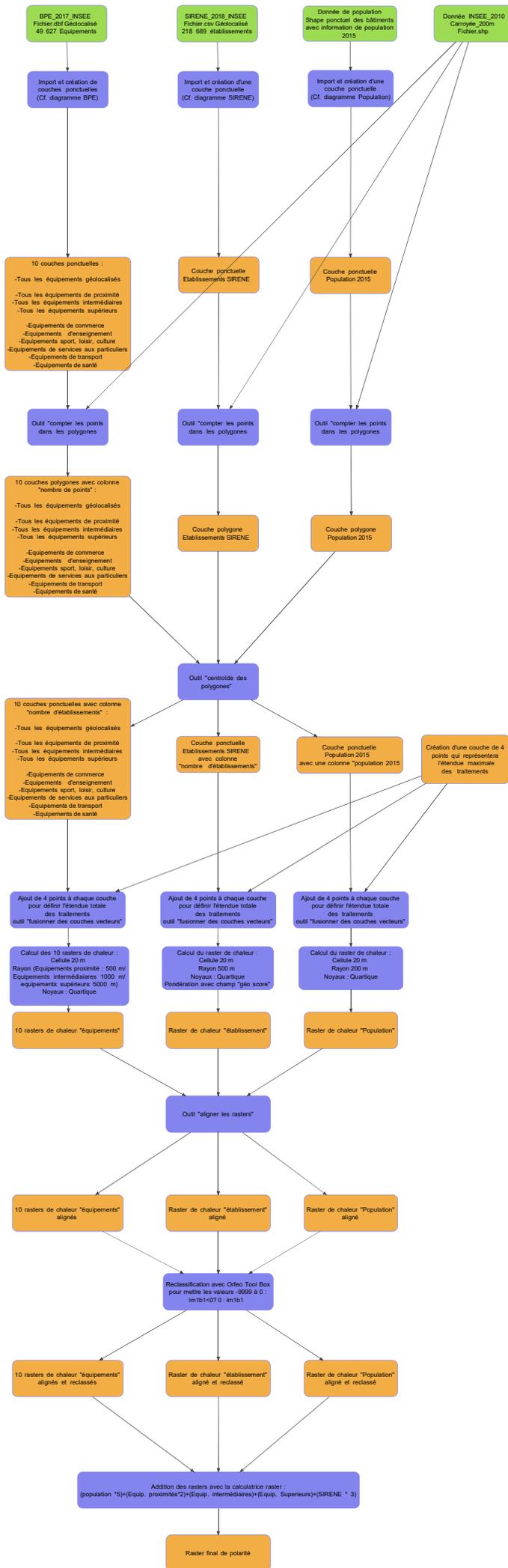
Traitement de la Base Permanente des Equipements de l'INSEE (BPE 2017) :



Légende :



Traitement des données par maille pour générer l'indice de polarité :



Légende :



Traitement des indicateurs par point :

