



Construction d'une base de données géographiques exhaustive à échelle fine sur l'occupation agricole du sol : le RPG complété

Benjamin Lardot, Pierre Cantelaube, Marie Carles, Claire Seard, Camille Truche

► To cite this version:

Benjamin Lardot, Pierre Cantelaube, Marie Carles, Claire Seard, Camille Truche. Construction d'une base de données géographiques exhaustive à échelle fine sur l'occupation agricole du sol : le RPG complété : Partie 1 : Production de la couche géographique des parcelles susceptibles d'accueillir les surfaces agricoles hors RPG.. 2022. hal-03647500

HAL Id: hal-03647500

<https://hal.inrae.fr/hal-03647500>

Submitted on 20 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



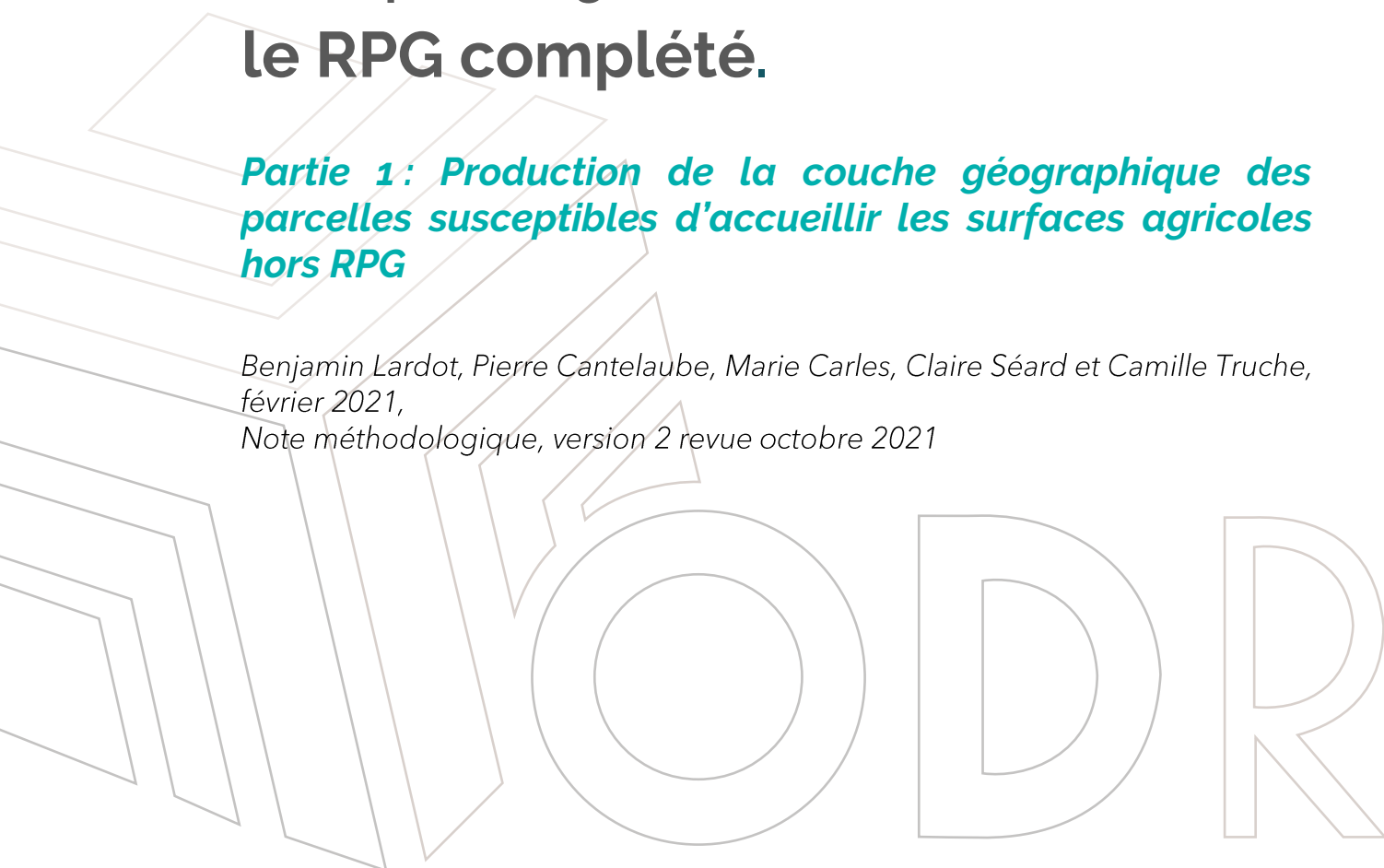
Construction d'une base de données géographiques exhaustive à échelle fine sur l'occupation agricole du sol :

le RPG complété.

Partie 1 : Production de la couche géographique des parcelles susceptibles d'accueillir les surfaces agricoles hors RPG

Benjamin Lardot, Pierre Cantelaube, Marie Carles, Claire Séard et Camille Truche, février 2021,

Note méthodologique, version 2 revue octobre 2021



La note méthodique décrivant la création et la construction du RPG complété est constituée de 3 documents :

1. **Production de la couche géographique des parcelles susceptibles d'accueillir les surfaces agricoles hors RPG**
2. Attribution des cultures aux parcelles identifiées
3. Résultats : présentation et analyses (campagne 2018)

Auteur de correspondance : pierre.cantelaube@inrae.fr

Pour citer ce document :

Lardot B., Cantelaube P., Carles M., Séard C. et Truche C., 2021, Construction d'une base de données géographiques exhaustive à échelle fine sur l'occupation agricole du sol : Le « RPG complété ». *Partie 1 : Production de la couche géographique des parcelles susceptibles d'accueillir les surfaces agricoles hors RPG*. Note Méthodologique US ODR, INRAE, février 2021, version 2 revue octobre 2021.

(1^{er} version : février 2021)

Version 2, octobre 2021.

<https://odr.inrae.fr/>

Les travaux présentés dans ce document ont été réalisés sous financements du projet OFB dans le cadre de la convention OFB-INRAE (2020-2021). Les travaux précédents ont pu être menés à bien grâce aux financements dans le cadre du Projet 11 du projet spatialisation BNVD- plan Ecophyto (2012-2016 puis 2017-2020)



Table des matières

Avant-propos et résumé.....	5
1. Introduction.....	6
2. Données.....	1
2.1- Corine Land Cover 2018	1
2.2- BD Parcellaire Express® 2020 V1 (IGN)	1
2.3- BD TOPO® 2019 V3.0 (IGN)	1
2.4- BD CARTO® 2018 V3.3 (IGN)	2
2.5- RPG 2015 à 2019 (IGN).....	2
2.6- Petite Région Agricole	2
2.7- Table de passage (INSEE)	2
2.8- OSO (CESBIO).....	2
3. Principe général de la procédure	3
4. Parallélisation et mémoire RAM	5
5. Librairies <i>Python</i>	5
6. Fonctions de nettoyage.....	5
6.1 Nettoyage des géométries	5
6.2 Nettoyage selon la forme.....	5
6.3 Nettoyage selon la surface	6
7. Préparation des <i>Parcelles ODR</i>	7
7.1 Préparation des données	8
7.2 Construction du <i>Masque 1</i>	9
7.3 Différence Parcelles - RPG (<i>Parcelles 1</i>)	10
7.4 Différence <i>Parcelles 1 – Masque 1 (Parcelles 2)</i>	10
7.5- Construction du <i>Masque 2</i> (Parcelles bâties).....	10
7.6 Différence <i>Parcelles 2 - Masque 2 (Parcelles 3)</i>	11
7.7 Intersection avec <i>CLC (Parcelles 4)</i>	11
7.8 Traitement des Vignes, Vergers (<i>Parcelles 5</i>).....	11
7.9- Traitement des Serres Agricoles (<i>Parcelles 6</i>).....	11
7.10 Multi polygones vers polygones (<i>Parcelles 7</i>)	12
7.11 Jointure avec les Communes PRA	12
7.12 Finalisation des Parcelles ODR	12
7.13 Calcul des statistiques zonales OSO	14
7.14 Calcul des statistiques zonales liées à l'altitude.....	15
8. Conclusion	17

9. Conclusion sur la phase 1 « <i>identification des parcelles agricoles hors RPG</i> » et perspectives	17
9. Références	19
10. Annexe - Notions de cadastre et de parcellaire agricole	20

Avant-propos et résumé

La cartographie d'occupation du sol agricole a été élaborée à l'INRAE, unité ODR, dans le cadre du programme EcoPhyto. Le but est de compléter le RPG pour disposer, à l'échelle parcellaire, d'une représentation géographique exhaustive de la surface agricole française (métropolitaine). Il s'agit donc de créer une carte annuelle parcellaire, nommée RPG complété dont l'objectif peut se résumer grâce à la formule suivante : *RPG + RPG complété = couverture exhaustive de la surface agricole* (La surface agricole étant définie comme l'ensemble de la superficie des terres arables, des prairies permanentes et des pâturages permanents ou des cultures permanentes, voir CCE, 2016)

La méthodologie développée permet donc de cartographier un ensemble de parcelles agricoles absentes du RPG, nommées ci-après les « parcelles ODR ». Se décomposant en 2 étapes, le travail consiste (1) à repérer ces parcelles susceptibles d'accueillir ces surfaces agricoles qui n'appartiennent pas à des exploitants recevant des aides PAC surfaciques, puis (2) à déterminer les cultures présentes sur ces parcelles identifiées.

De telles cartes (dénommées donc « RPG complété ») ont été établies pour compléter les campagnes 2012 à 2014 du RPG (ODR-Ecophyto, 2015, Cantelaube et al., 2016, Séard, 2016) suivant une première méthodologie (RPG complété version v1) revue et améliorée.

Initialement, dans cette première version du RPG complété, l'attribution des cultures aux parcelles reposait sur un traitement statistique et ciblait la différence entre les surfaces présentes dans le RPG et les statistiques départementales sur les surfaces agricoles fournies par le service Agreste du Ministère de l'Agriculture (Statistique Agricole Annuelle, SAA) (Agreste, 2014). Un modèle probabiliste permettait d'attribuer aux parcelles ODR une ou plusieurs catégories de cultures (nous parlions alors d'« îlot ODR » plutôt que de « parcelles ODR ») de façon à ce que la somme des surfaces par culture ainsi obtenue, sur un département, corresponde exactement à la différence SAA - RPG. Les résultats avaient du sens uniquement à une échelle agrégée (groupes de communes, cantons, petites régions agricoles, etc.) mais n'étaient pas garantis à l'échelle de la parcelle. De plus, les biais de comparaison entre le RPG et la SAA (surfaces des exploitations ayant leur siège social dans le département vs surfaces physiquement dans le département, différences temporelles, de nomenclatures) ainsi que le niveau géographique des entrées du modèle probabiliste (principalement l'échelle de la carte CORINE land COVER, CLC, 2009) rendaient ces résultats que modérément satisfaisants et peu utilisables à une échelle fine.

Dans la version 2 du RPG complété présentée ici, cette volonté de viser absolument la référence fournie par la SAA est abandonnée. S'en détacher est rendu possible grâce à un nouvel apport d'information thématique sur l'occupation du sol à une échelle fine : l'occupation du sol fournit par l'imagerie satellitaire via la carte OSO produite par le Centre d'Etudes Spatiales de la BIOshère (Cesbio). Cette information auxiliaire, issue de la télédétection, de qualité (thématiquement, géographiquement et temporellement) permet en outre d'attribuer une seule culture par parcelle et est devenue le support principal du RPG complété. Après avoir évalué l'utilisation d'OSO dans notre contexte et préparer notre méthode de construction à partir des cartes OSO pré-2018 (OSO en 17 classes, dont 2 agricoles + 1 classe dédiée aux prairies), une méthodologie plus « définitive » a été établie pour l'intégration de la carte OSO version « définitive » en 23 classes (dont 9 agricoles + 1 classe dédiée aux prairies).

L'étape de construction du fond parcellaire des surfaces potentiellement agricoles et non présentes dans le RPG a également été revue, notamment en intégrant les mises à jours faites sur les données d'entrées des différents traitements géomatiques et en encapsulant ces derniers dans un programme global écrit en langage de programmation Python.

1. Introduction

Le Registre Parcellaire Graphique (RPG) est produit chaque année par l'Agence de Services et de Paiement (ASP) en collaboration avec l'Institut National de l'Information Géographique et Forestière (IGN). Cette base de données géographique recense les parcelles agricoles françaises (IGN, 2020b, Cantelaube et Carles, 2015). Le RPG présente l'inconvénient de ne recenser que les parcelles des exploitations agricoles sujettes à des aides de la Politique Agricole Commune (PAC, aides du 1^{er} pilier et aides surfaciques du 2nd pilier). Une méthodologie a été développée par l'unité INRAE US ODR pour identifier et délimiter des parcelles dites « Parcelles ODR » susceptibles d'être ces parcelles agricoles absentes du RPG¹.

La combinaison du RPG et du RPG complété permet donc d'obtenir une couche exhaustive parcellaire et annuelle de la couverture du sol agricole français (hors DOM). Cette couche de données RPG complété est produite depuis 2014 pour couvrir les campagnes 2012 à 2014 (RPG complété version v1). La méthodologie a été revue pour produire le RPG complété version v2 pour les campagnes 2014 (refaite donc) et suivantes.

Le RPG complété est construit en 2 phases. **La 1^{ère} phase vise à identifier les parcelles susceptibles d'accueillir les parcelles agricoles absentes du RPG.** Cette étape s'effectue via une succession de traitements géomatiques grâce à l'utilisation de nombreuses couches d'informations géographiques, notamment issues de l'IGN : Bd Parcellaire® (Cadastre), Bd Topo®. **La 2^{nde} phase attribue à ces parcelles identifiées une catégorie d'occupation du sol** (culturales ou non agricole). Cette dernière étape utilise de nombreuses informations auxiliaires ; et si la version v1 se basait essentiellement sur la statistique agricole annuelle (SAA, Agreste) et la couche CORINE Land Cover (CLC), la version v2 utilise désormais la carte d'occupation du sol issue de la télédétection ; la carte OSO développée par le CESBIO-theia².

Le présent document décrit la méthodologie globale - revue en 2021 - pour produire **le fond parcellaire agricole hors RPG**, que l'on nomme donc **Phase 1** de l'élaboration du RPG complété. Une attention particulière a été portée sur l'intégration des évolutions des données d'entrées et l'automatisation des traitements afin de faciliter la production de ce fond de carte dès qu'un nouveau RPG est mise à disposition par ses producteurs, à partir de données auxiliaires les plus contemporaines possible avec la campagne du RPG visée.

Note : sont nommées « parcelles odr » les parcelles identifiées lors de la phase 1 auxquelles est attribuée une catégorie de culture lors de la phase 2.

¹ Ce travail a été réalisé dans le cadre du projet *Spatialisation de la Base Nationale des Ventes des produits phyto (BNV-d)* financé dans le cadre du plan Ecophyto (ODR-Ecophyto, 2015 ; Séard C., 2016 ; Cantelaube et al., 2016).

² <https://www.theia-land.fr/ceslist/ces-occupation-des-sols/>

2. Données

Les bases de données utilisées sont toutes diffusées sous licence ouverte. Elles proviennent de l'IGN (<http://www.ign.fr/>), de la base de données européenne *Corine Land Cover (CLC)* (<https://land.copernicus.eu/pan-european/corine-land-cover>) et du Centre d'Etudes Spatiales de la BIOsphère (CESBIO) (<https://www.cesbio.cnrs.fr/>).

Note : Toutes les données utilisent le système de projection *Lambert-93* sauf *CLC*. Cette couche est donc re-projetée en *Lambert 93* préalablement à tout traitement. Ainsi, l'ensemble des données est superposable. Les échelles de production des différentes données sont cohérentes et inférieures au 1:50 000 sauf pour *CLC* (1 : 100 000). A titre d'exemple, la surface de la plus petite unité cartographiée dans *CLC* (seuil de description) est de 25 hectares.

	<i>IGN</i>	<i>CLC</i>	<i>CESBIO (OSO)</i>
<i>Système de projection</i>	RGF93 / Lambert-93 (SCR EPSG : 2154)	Lambert II étendu	RGF93 / Lambert-93 (SCR EPSG : 2154)
<i>Format</i>	Shapefile	Shapefile	Tif
<i>Résolution spatiale</i>			10 x 10 m
<i>Echelle(s)</i>	1 : 2 500 au 1 : 50 000 selon les BD	1 : 100 000	
<i>Etendue</i>	France	Europe	France

2.1- Corine Land Cover 2018

CLC est un inventaire biophysique de l'occupation du sol et de son évolution selon une nomenclature en 44 postes. CLC est mis à jour tous les 6 ans. Elle est utilisée pour attribuer aux *Parcelles ODR* le code CLC d'occupation du sol correspondant (« code_18 »).

Documentation : [Guide d'utilisation CLC](#)

2.2- BD Parcellaire Express® 2020 V1 (IGN)

La *BD Parcellaire* « historique » n'est plus entretenue depuis son ultime édition en 2018. Elle est remplacée par la *BD Parcellaire Express®*, produit régulièrement mis à jour et plus exhaustif. La *BD Parcellaire Express®* est issue du PCI vecteur de la *DGFiP* et sa structure est analogue à celle de la *BD Parcellaire* (même couches et même tables attributaires), mais sans les corrections topologiques du processus *BD Parcellaire* ce qui peut entraîner parfois des problèmes de cohérence avec d'autres couches. De plus quelques départements ne sont pas (encore) intégralement couverts par cette BD (08,10,52,55,67,70).

Note : Pour les départements où la *BD Parcellaire Express®* est incomplète, la *BD Parcellaire* « historique » est utilisée si toutefois elle est plus exhaustive.

Couches : *PARCELLE* et *BATIMENT*.

Documentation : [DC DL SE Comparatif IGN - BD Parcellaire VS BD Parcellaire Express](#)

2.3- BD TOPO® 2019 V3.0 (IGN)

[La BD TOPO® est une description vectorielle tridimensionnelle](#) du territoire et de ses infrastructures de précision métrique. [La V3.0](#) change beaucoup de structure (thèmes, classes, attributs et valeurs d'attributs) par rapport aux versions précédemment utilisées.

La mise à jour des thèmes *BATI*, *OCCUPATION_DU_SOL* et *TRANSPORT* suit le cycle de mise à jour des prises de vues aériennes (de 3 à 4 ans). Le thème *ADMINISTRATIF* est mis à jour en continu, et fiabilisé annuellement à partir du code géographique de l'INSEE. Enfin, concernant la couche *ZONE_D_ACTIVITE_OU_D_INTERET* du thème *SERVICES_ET_ACTIVITE*, les éléments d'importance au moins départementale (établissements administratifs, sportifs, de santé, d'enseignement...) sont

actualisés tous les 1 à 2 ans à partir de diverses sources partenariales et consolidées par une expertise documentaire et un passage terrain.

La V3.0 est donc la plus susceptible de se rapprocher de la vérité terrain pour les années 2015 à 2019.

Thèmes : ADMINISTRATIF, OCCUPATION DU SOL, HYDROGRAPHIE, TRANSPORT, BATI, SERVICES ET ACTIVITE.

Couches : COMMUNES, DEPARTEMENT, ZONE VEGETATION, SURFACE HYDROGRAPHIQUE, TRONCONS DE VOIE FERREE, TRONCON DE ROUTE, AERODROME, EQUIPEMENT DE TRANSPORT, BATIMENT, TERRAIN DE SPORT, CIMETIERE, ZONE D'ACTIVITE ET D'INTERET.

Documentation : [Docs IGN](#)

2.4- BD CARTO® 2018 V3.3 (IGN)

La BD CARTO® est une représentation vectorielle bidimensionnelle du territoire et de ses infrastructures. La couche **ZONE_OCCUPATION_SOL** utilisée est élaborée à partir d'images satellitaires de 2018. Elle permet d'accéder à certaines occupations du sol que l'on ne retrouve pas dans la BD TOPO® (Carrières, décharges, névés, glaciers, marais, rochers, broussailles ou sable).

Documentation : [Docs IGN](#)

2.5- RPG 2015 à 2019 (IGN)

Le *Registre Parcellaire Graphique (RPG)* est un système d'information géographique géré par l'Agence de Service et de Paiement (ASP) et l'IGN permettant l'identification des parcelles agricoles. Les données du RPG fournissent des informations détaillées sur l'occupation du sol et la localisation des parcelles des exploitants agricoles ayant déposé une déclaration en vue d'une aide de la PAC. Les données sont déclarées par les exploitants et mises à jour annuellement.

Documentation : [Docs IGN](#)

2.6- Petite Région Agricole

Les Petites Régions Agricoles (PRA) sont des régions ayant une même vocation agricole dominante ; elles ont été délimitées par l'INSEE en 1946. La dernière actualisation date de 1981. Le *Shapefile* utilisé (*Communes PRA*) attribut à chaque commune un code PRA.

2.7- Table de passage (INSEE)

Cette table, diffusée par l'INSEE permet d'attribuer aux communes de la donnée « *Communes PRA* » un code INSEE actualisé (2020).

2.8- OSO (CESBIO)

Le CESBIO produit annuellement une carte d'occupation des sols à partir d'imagerie satellitaire (séries d'images optiques multi-temporelles à haute résolution spatiale). Le raster est fourni à l'échelle nationale avec une résolution spatiale de 10 mètres et une nomenclature à 23 postes (depuis 2018, 17 avant). Cette donnée nous permet de caractériser les parcelles détectées en vue de faciliter par la suite l'attribution des cultures. Le raster utilisé est préalablement découpé par emprise départementale.

Documentation : [OSO](#)

3. Principe général de la procédure

Un programme principal (*prog_principale_France_cfg.py*) appelle un sous-programme qui contient les fonctions nécessaires (*sous_prog_fct_cfg.py*). Deux variables sont nécessaires en entrée : la liste des codes *INSEE* des départements métropolitains à traiter (*numDep*) et l'année du *RPG* à considérer (*annee*).

Un *fichier de configuration* (*fichier_de_config.txt*), associé aux deux scripts *Python*, permet de lancer facilement le programme. C'est dans ce fichier texte que l'utilisateur saisit les chemins d'accès aux données, les dossiers d'enregistrement des résultats, les départements à traiter ainsi que l'année à prendre en compte (*RPG*), assurant la portabilité du programme.

Un *fichier de Log* est créé en début de programme permettant de suivre l'avancement du programme en temps réel.

Le programme principal se compose de deux parties :

► **Partie 1** - Traitement des Parcelles ODR.

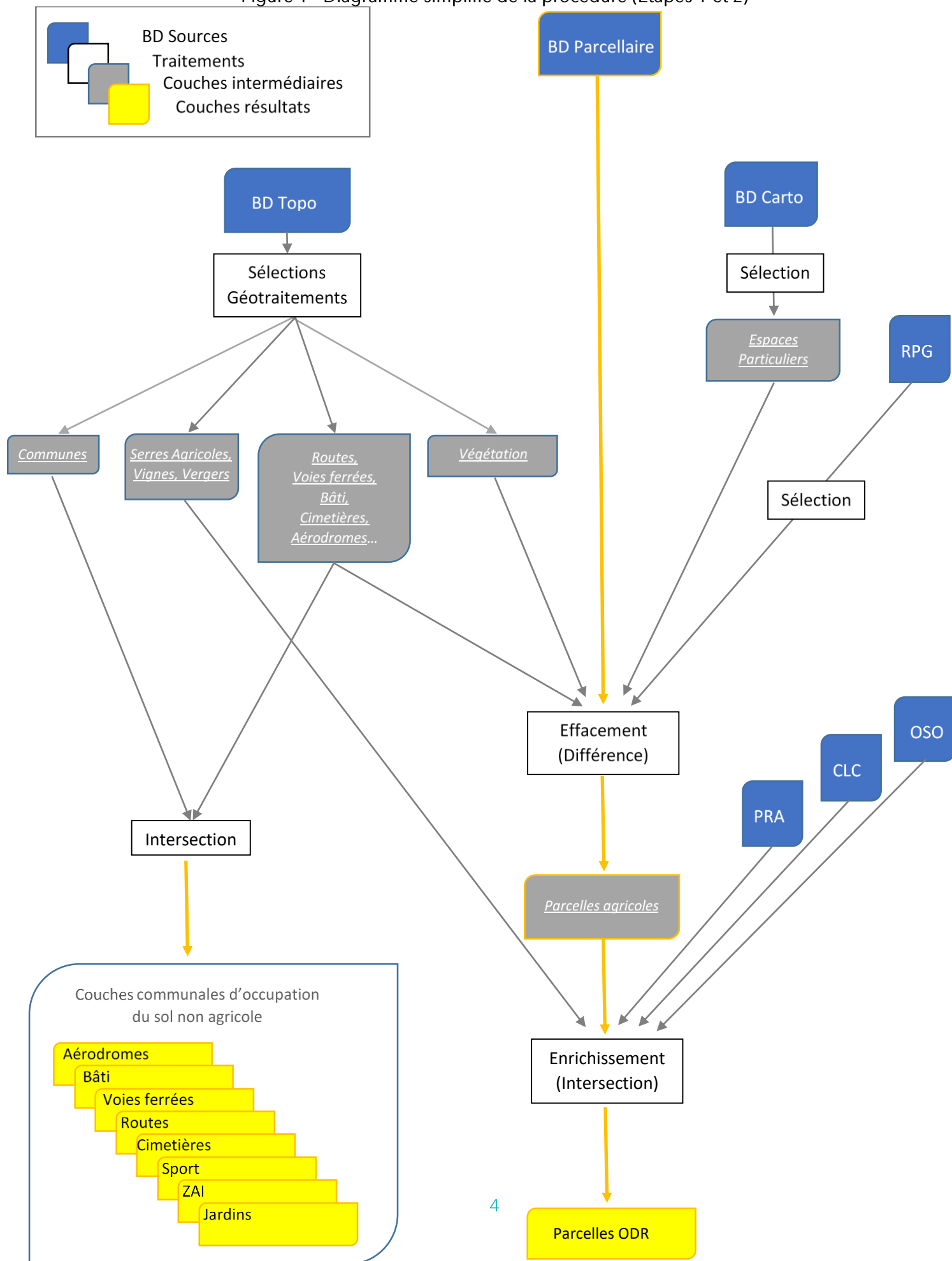
Ces parcelles sont construites à partir du parcellaire cadastral de la *BD Parcellaire Express*®. Les différentes étapes du traitement viennent découper et supprimer les parcelles cadastrales déjà présentes dans le *RPG* ou identifiées comme étant non agricoles (surfaces bâties, végétation, routes...). Les parcelles ou morceaux de parcelles restants sont ensuite intersectées avec *CLC* puis avec les Vignes, les Vergers, les Serres (issus de la *BD TOPO*®), les communes, les PRA et enfin la donnée *OSO* pour en récupérer les contours et/ou l'information.

► **Partie 2** - Traitement des couches d'occupations du sol non agricole susceptibles de recevoir des traitements phytosanitaires.

8 couches d'occupation du sol spécifiques sont obtenues par sélection et géo-traitement des bases de données sources et valorisent les traitements de la première partie.

Citée ici pour information, cette partie n'est pas présentée dans ce document

Figure 1 - Diagramme simplifié de la procédure (Etapes 1 et 2)



4. Parallélisation et mémoire RAM

Afin d'accélérer la production, le programme est parallélisé (module « *Multiprocessing* »), permettant de répartir les traitements sur plusieurs CPU. Plusieurs départements peuvent ainsi être traités simultanément.

Le facteur limitant est la mémoire vive de la machine utilisée. Le script est assez « lourd », nécessitant de stocker de grandes quantités de données. D'autre part certaines opérations géométriques sont très coûteuses en ressources (différences et intersections).

Pour limiter l'utilisation de la mémoire vive, les données sont effacées à mesure qu'elles deviennent obsolètes par le module « *Garbage Collector* » qui revient régulièrement dans le script. Pour compléter ce processus de nettoyage, certaines données sont placées à « None » dès lors qu'elles ne sont plus utiles.

La machine utilisée pour les campagnes 2015 à 2018 dispose de 16 Go de RAM et de 8 cœurs permettant de traiter 5 départements simultanément sans problème. Le temps moyen de traitement par département est de 2 heures (avec parallélisation), soit un peu plus de 7 jours pour l'ensemble des départements métropolitains.

5. Bibliothèques Python

- *Pandas* et son extension *geopandas* permet de manipuler aisément de grands volumes de données avec une composante spatiale : lecture, jointure (spatiale et attributaire), concaténation et opérations spatiales.
- *Time* et *Datetime* : pour mesurer les temps de traitement.
- *Glob* : pour faciliter la recherche des chemins d'accès aux données en entrée.
- *Multiprocessing* : permet la parallélisation en répartissant les traitements sur plusieurs processeurs.
- *Math* pour quelques calculs mathématiques.
- *Logging* pour créer et remplir le fichier de log.
- *Garbage collector* pour nettoyer la mémoire.
- *Rasterstats* pour calculer les statistiques zonales.
- *Numpy*.

6. Fonctions de nettoyage

La généralisation des traitements à l'échelle nationale produit parfois des polygones indésirables : trop petits, géométries invalides ou vides, polygones très fins et longs (bordure de routes, de parcelles). Pour anticiper d'éventuels problèmes et en vue d'alléger les données à manipuler, 3 types de nettoyage sont appliqués après les grandes opérations spatiales du script :

6.1 Nettoyage des géométries

- **Buffer 0** sur la colonne géométrie :
`parcelles["geometry"] = parcelles.buffer(0)`
- On écarte les **géométries vides** :
`parcelles = parcelles.loc[~parcelles.geometry.is_empty]`
- On écarte également les rares entités avec des **géométries invalides** :
`parcelles = parcelles[(parcelles['geometry'].is_valid)]`

6.2 Nettoyage selon la forme

L'indice de circularité de *Miller* est utilisé pour écarter certains polygones particuliers :

$$\text{Indice de Miller} = (4 * \pi * \text{Area}) / \text{périmètre}^2$$

Il s'agit du rapport de la superficie d'une entité, à celle d'un cercle de même périmètre. Cette indice varie de 0 à 1. Il est de 0 pour une ligne et de 1 pour un cercle.

Il permet d'écarter les polygones très fins et longs (résidu fréquent après les opérations de différences).

Le choix a été fait d'éliminer les polygones dont l'indice de circularité < 0.015 après les trois principales opérations spatiales du script. Les tests effectués ont montré que ce seuil (très bas) éliminait uniquement des polygones indésirables s'apparentant à des lignes.

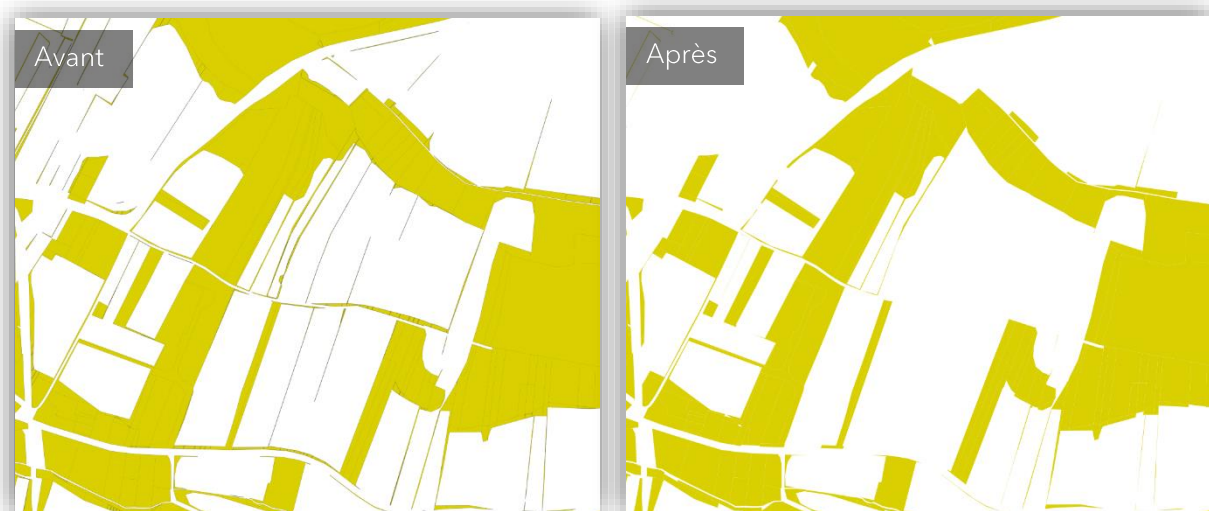


Figure 2 - Nettoyage avec l'indice de Miller (< 0.015)

Note : l'indice de Miller est également utilisé pour le nettoyage final des *Parcelles ODR* avec un autre seuil (0.12).

6.3 Nettoyage selon la surface

On calcule la surface des entités puis on élimine celles inférieures à 50 m^2 . Ces parcelles seront de toute façon aussi éliminées en fin de procédure ([7.12 Finalisation des Parcelles ODR](#)).

7. Préparation des *Parcelles ODR*

Cette partie présente les différentes étapes du programme par ordre chronologique.

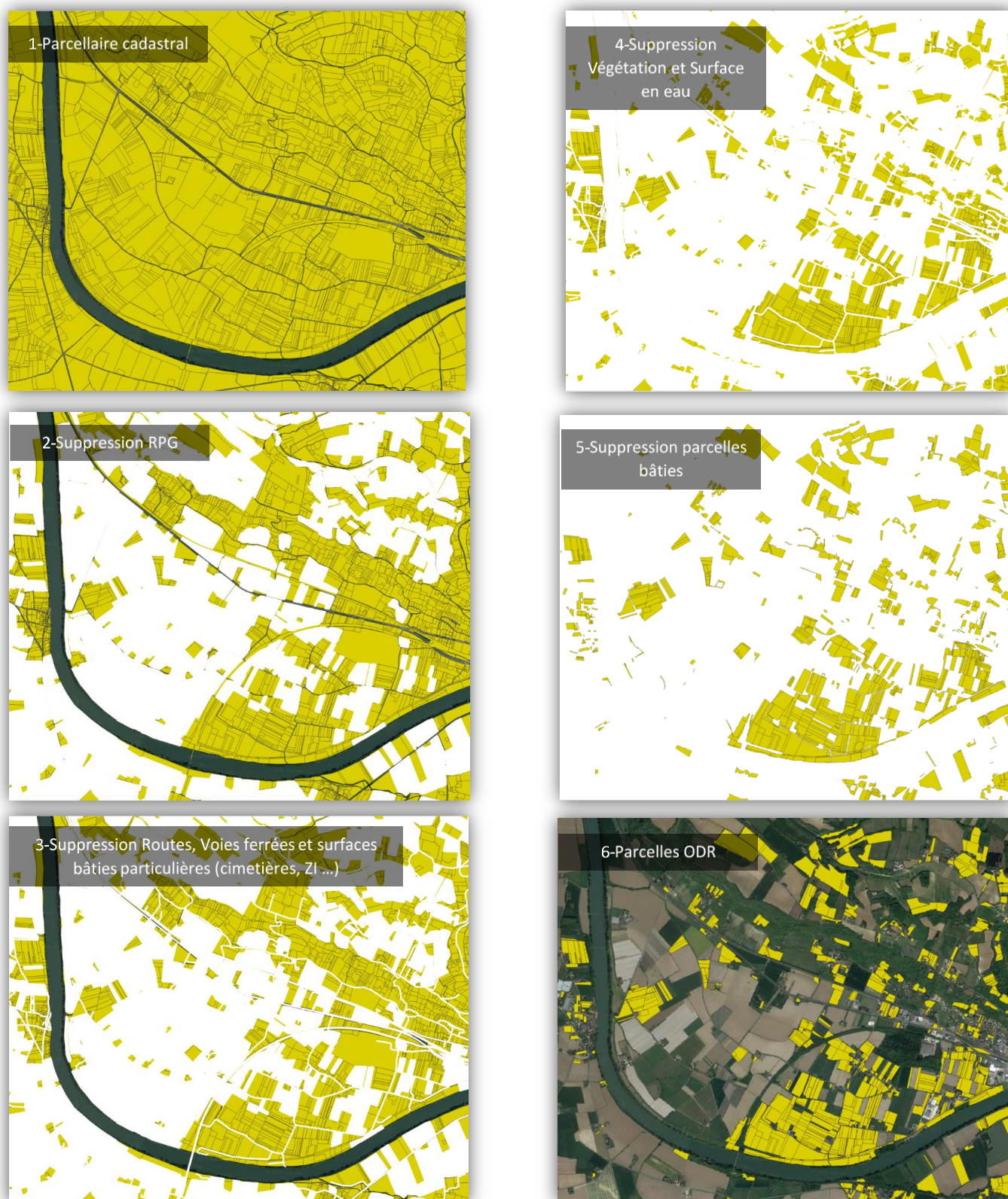


Figure 3 - Premières étapes du processus de production des Parcelles ODR (Lot-et-Garonne-47)

7.1 Préparation des données

Les principales données nécessaires sont préparées ici. Il s'agit de sélectionner les données relatives au département à traiter et de mettre en forme leurs tables attributaires. Sept fonctions sont mobilisées pour ce travail.

Note : Les fonctions utilisées sont indiquées en vert et les couches *Shapefile* nécessaires en orange. Les noms des attributs utilisés sont indiqués entre guillemets doubles (ex. « NATURE ») et les valeurs d'attribut entre guillemets simples (ex. 'Agricole').

➤ Préparation du département traité :

`def geomDep(numDep)` - BD TOPO - *DEPARTEMENT*

Sélection du département à traiter.

➤ Préparation des communes :

`def geomCom(numDep)` - BD TOPO - *COMMUNE*

Sélection des communes du département. Les communes sont utilisées dans la deuxième partie du script pour produire les 8 couches d'occupation du sol qui sont à l'échelle communale.

➤ Préparation de CLC :

`def clc_decoup(numDep)` - CLC - *CLC 2018*

Découpage de CLC sur l'emprise du département. Le résultat de cette fonction permettra d'ajouter aux *Parcelles ODR* le code CLC d'occupation du sol par intersection.

➤ Préparation des parcelles cadastrales :

`def parcelles(numDep)` - BD PARCELLAIRE EXPRESS - *PARCELLE*

- Sélection des parcelles cadastrales du département.
- Calcul du champ « *surfinitm2* » (surface initiale de la parcelle avant les découpages).
- Création du champ « *CODE_INSEE* » : concaténation des colonnes « *CODE_DEP* » et « *CODE_COM* ».
- Création du champ « *num_parce* » (identification de la parcelle) : concaténation des colonnes « *CODE_DEP* », « *CODE_COM* », « *FEUILLE* », « *SECTION* », et « *NUMERO* ».

➤ Préparation du RPG :

`def rpg(numDep,annee)` - *RPG*

Sélection des parcelles *RPG* pour le département concerné et l'année étudiée.

Note : Cette fonction sélectionne non seulement les parcelles *RPG* du département traité mais aussi les parcelles *RPG* des départements limitrophes. Cela permet d'éviter des erreurs liées à la préparation de la donnée *RPG* par département en amont (certaines parcelles en limite de communes ne sont pas attribuées au « bon » département).

➤ Préparation des serres agricoles :

`def select_serres(numDep)` - BD TOPO - *BATIMENT*

Sélection sur les champs « *NATURE* » = 'Serre' et « *USAGE1* » = 'Agricole'.

Les serres seront intersectées avec les *Parcelles ODR* pour en conserver les contours et l'information.

➤ Préparation des vignes et vergers :

`def select_vignes(numDep)` - BD TOPO - *ZONE_DE_VEGETATION*

Sélection sur le champ « NATURE » = 'vignes', 'vergers', 'bananeraie', 'canne à sucre', 'houblonnière' et 'rizière'.

Les Vignes et Vergers seront intersectées avec les *Parcelles ODR* pour en conserver les contours et l'information.

7.2 Construction du Masque 1

Quatre fonctions successives permettent de sélectionner des territoires (à partir des *BD TOPO* et *CARTO*) dont on sait qu'ils ne sont pas à vocation agricole. Les objets sélectionnés par ces fonctions sont alors regroupés pour former le *Masque 1*. Celui-ci sera ensuite retranché du parcellaire cadastral par différence géométrique.

➤ Voies ferrées :

`def voies_ferrees (numDep)` - BD TOPO - *TRONCON_DE_VOIE_FERREE*

La géométrie linéaire de cette donnée impose de la transformer en polygones. Nous estimons qu'une voie de chemin de fer fait 5 mètres de large. On applique donc un buffer de (2,5 mètres x « NB_VOIES ») aux tronçons sauf ceux 'En projet'. (L'attribut « NB_VOIES » est obligatoire dans la *BD TOPO* et est donc toujours renseigné).

➤ Les routes :

`def routes (numDep)` - BD TOPO - *TRONCON_DE_ROUTE*

De même, la géométrie linéaire de cette donnée impose de la transformer en polygones.

(Sélection des tronçons 'Autoroute', 'Départementale', 'Nationale', 'Route européenne', 'Route intercommunale' (attribut « CL_ADMIN »). Cette étape est supprimée. Cela permet de prendre en compte tous les tronçons de route pour la construction du masque et donc d'affiner les résultats.

Sélection de tous les tronçons de route dont l'attribut « NATURE » est différent de 'Escaliers' et de 'Bac ou liaison maritime'.

Ces tronçons sont bufferisés selon un champ « demilarg » créé et rempli à partir de l'attribut « LARGEUR ». Lorsque le champ « LARGEUR » n'est pas renseigné, on attribue au champ « demilarg » les valeurs suivantes : 'Type autoroutier':10, 'Bretelle' :1.5, 'Route à 2 chaussées':1.5, 'Route à 1 chaussée':1.5, 'Route empierrée':1.5, 'Chemin':1, 'Rond-point':2.5, 'Piste cyclable':0.5, 'Sentier':0.5. Les largeurs utilisées ici sont tirées de la note méthodologique de Seard C. (septembre 2016) « *Mise à jour de la cartographie d'occupation du sol* ».

Note : Dans la *BD TOPO*®, les routes à une chaussée sont matérialisées par une ligne et celles à deux chaussées par deux ; on leur attribue donc la même largeur (1.5 m).

➤ Une sélection de bâti de la *BD TOPO*® :

`def select_bati1 (numDep)`

Cette fonction sélectionne les données dans différentes couches de la *BD TOPO*® et les concatène.

- *BD TOPO* - *CIMETIERE* : Tous les cimetières.
- *BD TOPO* - *TERRAIN_DE_SPORT* : Tous les terrains de sport.
- *BD TOPO* - *AERODROME* : Tous les aérodromes dont la géométrie est non fictive (attribut « FICTIF » = 'Non').
- *BD TOPO* - *ZONE_D_ACTIVITE_OU_D_INTERET* : Toutes celles dont l'attribut « NATURE » n'est pas 'Divers agricole' et dont la géométrie est non fictive (attribut « FICTIF » = 'Non').

- BD TOPO - **EQUIPEMENT_DE_TRANSPORT** : Tous ceux dont la géométrie est non fictive (parkings, aires de triage, aires de repos...).
- Une sélection de non bâti issue des BD TOPO® et CARTO® :
`def select_non_bati1 (numDep)`
 - BD TOPO - **ZONE_DE_VEGETATION** : Sélection des 'Bois', 'Forêt fermée de conifère', 'Forêt fermée de feuillus', 'Forêt fermée mixte', 'Forêt ouverte', 'Haie', 'Lande herbacée' (présent seulement en Guyane), 'Lande ligneuse', 'Mangrove', 'Peupleraie', et 'Zone arborée' du champ « NATURE ».
Note : La couche **ZONE_DE_VEGETATION** de la BD TOPO® est désormais « multi-thème » pour l'ensemble des départements métropolitain, c'est-à-dire que tous les types de couverts végétaux sont différenciés. Le « mono-thème (Attribut « NATURE » = 'Zone arborée') ne concerne plus aucun département. D'autre part, 'Lande Ligneuse' peut avoir une utilisation agricole mais nous considérons qu'il s'agit d'un biais mineur.
 - BD TOPO - **SURFACE_HYDROGRAPHIQUE** : Tout.
 - BD CARTO - **ZONE_OCCUPATION_SOL** : « NATURE » = 'Carrière - décharge', 'Glacier - névé', 'Marais', 'Rocher - éboulis', 'Sable - gravier', 'Mangrove', 'Broussailles'.

7.3 Différence Parcelles - RPG (Parcelles 1)

On retire des parcelles cadastrales celles incluses dans le RPG de l'année de référence par différence géométrique ; les parcelles résultantes sont nommées *Parcelles 1*.

`def overlay (parcelles, rpg, how="difference")`

Puis 3 nettoyages successifs :

- Nettoyage des géométries.
- Nettoyage selon la forme (*Miller* < 0.015).
- Nettoyage selon la taille (>50m²).

7.4 Différence Parcelles 1 – Masque 1 (Parcelles 2)

On retire ensuite des *Parcelles 1* les territoires compris dans le *Masque 1* par différence géométrique ; les parcelles résultantes sont nommées *Parcelles 2*.

`def overlay (parcelles1, masque1, how="difference")`

Puis 3 nettoyages successifs :

- Nettoyage des géométries.
- Nettoyage selon la forme (*Miller* < 0.015).
- Nettoyage selon la taille (>50m²).

Note : L'identifiant des parcelles construit lors de la préparation des données (« num_parc ») ne s'avère pas unique. On ajoute donc à cette étape un identifiant « vraiment » unique aux *Parcelles 2* (« id_parc » = « num_parc » + incrémentation). Il sera nécessaire pour la suite.

7.5- Construction du Masque 2 (Parcelles bâties)

Le *Masque 2* est une sélection des parcelles cadastrales qui contiennent du bâti. Ces parcelles n'ont pas de vocation agricole et seront donc retirées.

Deux précautions s'imposent :

- 1) Ne pas écarter les parcelles dont les bâtiments sont des Serres Agricoles.

Pour cela, les bâtiments (BD PARCELLAIRE EXPRESS - **BATIMENT**) qui intersectent des bâtiments de « NATURE » 'serre' et à « USAGE » 'agricole' (BD TOPO - **BATIMENT**) sont préalablement sélectionnés et écartés (ainsi ils ne participeront pas à la construction du *Masque 2*).

- 2) Ne pas écarter les parcelles contenant une petite surface de bâtiments par rapport à la surface totale de la parcelle (cas du petit bâtiment au milieu d'un champ de vigne...).

Pour cela, on calcul la surface totale de bâti par parcelle : les bâtiments de la *BD PARCELLAIRE* (tous, sauf ceux qui sont des serres donc) sont intersectés avec les *Parcelles 2*. La surface de chaque bâtiment est alors calculée puis sommée par parcelle. Seules les parcelles dont le ratio (surface de bâti / surface de parcelle restante) > 0,025 sont retenues dans le *Masque 2*. Ce seuil est obtenu empiriquement après de nombreux tests.

Nettoyage :

- Nettoyage des géométries.

7.6 Différence Parcelles 2 - Masque 2 (Parcelles 3)

On retire des *Parcelles 2* les territoires compris dans le *Masque 2* par différence géométrique; les parcelles résultantes sont nommées *Parcelles 3*.

```
def overlay (parcelles2, masque2, how="difference")
```

Puis 3 nettoyages successifs :

- Nettoyage des géométries.
- Nettoyage selon la forme (*Miller* < 0.015).
- Nettoyage selon la taille (>50m²).

Note : A cette étape, on calcul les surfaces restantes de chaque parcelle (« surfm2 »). Cette surface sera nécessaire pour la fonction de nettoyage final des résidus ([7.12 Finalisation des Parcelles ODR](#)).

7.7 Intersection avec CLC (Parcelles 4)

Les *Parcelles 3* sont intersectées avec *CLC*. Cela permet de les découper selon l'occupation du sol *CLC* et de capitaliser le « code_18 » et l'« id » (renommé ensuite « id_clc »).

```
def overlay (parcelles3, clc_dpt, how="intersection")
```

Nettoyage :

- Nettoyage des géométries.

7.8 Traitement des Vignes, Vergers (Parcelles 5)

Les *Parcelles 4* sont intersectées avec les *Vignes - Vergers* (donnée préparée en 6.1) pour en récupérer les contours et l'information. Les parcelles résultantes sont nommées *Parcelles 5* et sont enrichies d'un champ « NATURE1 » avec pour valeurs possibles 'vigne', 'verger', 'bananeraie', 'canne à sucre', 'rizière', 'houblonnière' ou 'NULL'.

Note : Le champ « NATURE1 » a la valeur 'NULL' quand l'intersection n'a pas donné de résultat.

Note : La couche *ZONE_DE_VEGETATION* de la *BD TOPO* comporte quelques erreurs pour les « NATURE » = 'vignes' et 'vergers'; ainsi des polygones de 'vignes' se superposent parfois à des polygones 'vergers' dédoublant les *Parcelles ODR* lors de cette intersection (une centaine de parcelles ODR sont dédoublées pour le millésime 2015). Ces doublons seront éliminés en post traitement via PostgreSQL en conservant les deux informations (Vignes et Vergers).

L'intersection se déroule en 3 étapes afin de conserver tous les enregistrements :

- Intersection (*Parcelles 4 - Vignes,Vergers*) = *Parcelles 5a*
- Différence (*Parcelles 4 - Vignes,Vergers*) = *Parcelles 5b*
- Concaténation (*Parcelles 5a + Parcelles 5b*) = *Parcelles 5*

7.9- Traitement des Serres Agricoles (Parcelles 6)

Comme pour les Vignes et Vergers, les *parcelles 5* sont intersectées avec les *Serres* (donnée préparée en 6.1) pour en récupérer les contours et l'information. Les parcelles résultantes sont nommées *Parcelles 6* et sont enrichies d'un champ « NATURE2 » avec pour valeurs possibles 'serre'

ou 'NULL'.

L'intersection se déroule en 3 étapes afin de conserver tous les enregistrements :

- Intersection (*Parcelles 5 - Vignes, Vergers*) = *Parcelles 6a*
- Différence (*Parcelles 5 - Vignes, Vergers*) = *Parcelles 6b*
- Concaténation (*Parcelles 6a + Parcelles 6b*) = *Parcelles 6*

Nettoyage :

- Nettoyage des géométries.

7.10 Multi polygones vers polygones (*Parcelles 7*)

Avec les géo-traitements successifs, la géométrie de certaines parcelles est devenue « multi parties ». La fonction « explode » permet de séparer les différents « morceaux » composant une parcelle. Les parcelles résultantes sont nommées *Parcelles 7*.

`def explode (parcelles6)`

7.11 Jointure avec les Communes PRA

Une fonction découpe préalablement la donnée *Communes PRA* sur l'étendue du département traité :

`def pra_decoup (numDep).`

Certains codes *INSEE* des communes PRA étant obsolètes, nous passons par une table de passage de l'*INSEE* pour attribuer à chaque commune PRA (et chaque *Parcelles 7*) le code géographique de l'année 2020 (« *CODGEO_2020* »).

Une jointure attributaire des *Communes PRA* sur les *Parcelles 7* via ce champ permet alors d'attribuer aux parcelles le « code PRA » correspondant.

7.12 Finalisation des Parcelles ODR

- Élimination des résidus 1 : `def Nettoyage (fil, how='residu')`

Les découpages successifs créent des résidus de parcelle dont la taille peut être importante mais qu'il est nécessaire d'éliminer du résultat. Ne travailler que sur la surface de la parcelle ne suffit pas à éliminer ces résidus. Ainsi, on travaille sur une combinaison entre la surface finale et le rapport entre surface après les découpages et surface initiale.

Cette fonction effectue une sélection des *Parcelles ODR* qui respectent la formule :

`((surffinale > 1500) & ((surfm2/ surfinitm2)>0.4) & ((surfm2/ surfinitm2) <0.8)) or ((surffinale>100) & ((surfm2/ surfinitm2)>0.7))`

Avec :

- « **surfinitm2** » : surface initiale de la parcelle en m².
- « **surfm2** » : surface de la parcelle en m² après le découpage des masques (*RPG*, végétation, surface bâtie...) mais avant la séparation des polygones (multi polygones vers polygones).
- « **surffinale** » : surface finale en m² de chaque polygone.

Note : Les bornes utilisées ici sont tirées de la note méthodologique de Seard C. (septembre 2016) « *Mise à jour de la cartographie d'occupation du sol* ».

Par ailleurs, cette fonction conserve tous les « morceaux » de serres, vignes ou vergers (qui peuvent être de petites dimensions) dès lors qu'ils font plus de 50 m² (contre 100 m² pour les autres) :

`(NATURE=='Verger')&(surffinale>50)|(NATURE=='Vigne')&(surffinale>50)|(NATURE=='Bananeraie')&(surffinale>50)|(NATURE=='Canne à sucre')&(surffinale>50)|(NATURE=='Rizi  re')&(surffinale>50)| (NATURE =='Houblonni  re')&(surffinale > 50)]`

- Élimination des résidus 2: `def Nettoyage (fil, how='miller')`

Le choix a été fait d'éliminer les objets dont l'indice de circularité (*Miller*) est inférieur à 0.12. Nous avons conservé la valeur utilisée dans le *model builder* initial.

- Calcul de la surface finale des Parcelles ODR: « surffinale ».
- Ajout d'un identifiant unique des Parcelles ODR: « id_fin ».
- Calcul de l'indice de circularité de Miller des Parcelles ODR: « Ind_circ »



surfinitm2	CODE_INSEE	code_18	id_clc	NATURE1	NATURE2 ▼	code_pra	surffinale	id_fin	Ind_circ
2395,1335...	90053	231	EU-475619		Serre	90303	223,3791296...	90_053_1...	0,76249...
2395,1335...	90053	231	EU-475619		Serre	90303	460,1749999...	90_053_1...	0,74632...
11440,252...	90067	242	EU-483660	Verger		90303	11209,19505...	90_067_1...	0,37008...



Figure 4 - Parcelles ODR détectés dans le Lot-et-Garonne (47) et table attributaire associée

7.13 Calcul des statistiques zonales OSO

Les *Parcelles ODR* sont maintenant enrichies de l'information OSO. Le module *Python Rasterstats* permet de calculer les statistiques zonales pour chaque *Parcelle ODR* à partir du raster OSO.

Figure 5 - Parcelles ODR (délimitées en rouge) superposées aux pixels OSO

Ainsi, 23 champs sont créés (« cl1 », « cl2 » ... « cl23 ») correspondant aux 23 classes d'occupation du sol de OSO (ou 17 champs pour les 17 classes d'occupation du sol avant 2018). Pour chacune d'elle le nombre de pixels intersectant la parcelle est compté. *Rasterstats* calcul également le nombre total de pixel des parcelles (« count_oso »), la valeur de pixel la plus petite (« min »), les valeurs de pixel en minorité et en majorité (« minority » et « majority »), la déviation standard (« std ») ainsi que la valeur médiane (« median »).

Note : c'est le centroïde de chaque pixel (et non toute la surface) qui est utilisé pour l'intersection avec les parcelles. Cela permet de limiter l'effet de bord.

Note : la nomenclature OSO ayant évolué entre 2017 et 2018 (passage de 17 postes à 23 postes), le script s'adapte automatiquement en fonction du millésime traité.

Ensuite, la variété (« variety ») est calculée en comptant le nombre de classes différentes présentes dans chaque parcelle.

Enfin 9 champs de regroupement sont créés à la faveur de la nomenclature emboîtée d'OSO :

- « **urbain** » = Urbain dense + Urbain diffus + Zones industrielles et commerciales + Surfaces routes
- « **cultures** » = Colza + Céréales à pailles + Protéagineux + Soja + Tournesol + Maïs + Riz + Tubercules/Racines
- « **prairies** » = Prairies
- « **vergers** » = Vergers
- « **vignes** » = Vignes
- « **forets** » = Forêts de conifères + Forêts feuillus
- « **pelouse** » = Pelouse
- « **landes ligneuses** » = Landes ligneuses
- « **divers** » = Surfaces minérales + Plages et dunes + Glaciers et neiges + Eau + Valeurs nulles

cl1	cl2	cl3	cl4	cl5	cl6 ▲	cl7	cl8	cl9	cl10	cl11	cl12	cl13	cl14	cl15	cl16	cl17	cl18	cl19	cl20	cl21	cl22	cl23
0	0	0	0	0	0	0	0	0	0	0	0	667	0	0	83	0	3	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	30	0	0	82	0	0	0	0	0	0	0
0	3	0	0	0	0	0	0	0	0	0	0	41	0	0	69	11	24	1	0	0	0	0

count	min	minority	majority	std	median	variety	urbain ▲	cultures	prairies	vergers	vignes	forets	pelouses	landes	divers
6	0	0	13	6,128...	13	1	0	0	4	0	0	0	0	0	2
3	5	18	5	6,128...	5	2	0	2	0	0	0	0	1	0	0
4	6	6	6	6,000...	12	2	0	2	0	0	0	0	2	0	0

Figure 6 - Champs correspondants aux statistiques zonales OSO ajoutés à la table attributaire des Parcelles ODR

Les *Parcelles ODR* sont alors exportées au format *Shapefile* en intégrant dans le nom des fichiers le numéro du département ainsi que l'année du RPG pris en compte (Ex. 01_Parcels_ODR_rpg_2017.shp).

7.14 Calcul des statistiques zonales liées à l'altitude

Les *Parcelles ODR* produites peuvent ensuite être enrichies d'information liées au relief (altitude, pentes, exposition) via une procédure *Python* également basée sur le module *Rasterstats*. Cette procédure (*zonal_stats_fct_alti_pente_odr.py*) n'est pas intégrée au script principal et nécessite de préparer préalablement des rasters départementaux de pentes, d'altitude et d'exposition à partir d'un Modèle Numérique de Terrain (MNT).

Pour les millésimes 2015 à 2019, c'est La *BD ALTI*® de l'IGN avec une résolution spatiale de 25 mètres qui est utilisée.

➤ Etape 1 : Préparation des rasters départementaux

La *BD Alti*® est tout d'abord découpée par département avec l'outil « Découper un raster selon une couche de masque » de *QGIS* en processus de lot. On calcul alors pour chaque département les pentes et les expositions avec les outils « pente » et « exposition » de *QGIS* en processus de lot. Nous obtenons donc pour chaque département trois rasters : altitude, pentes et exposition.

Note : l'altitude est exprimée en mètre et les pentes en degrés. Pour l'exposition, les valeurs sont comprises entre 0 et 360° et expriment la direction de la pente : à partir du nord (0°) et dans le sens horaire.

➤ Etape 2 : Procédure *Python*

La fonction *zonal_stats_fct_alti_pente_odr.py* est parallélisée pour traiter les *Parcelles ODR* de plusieurs départements simultanément et utilise un fichier de configuration (*fichier_de_config_odr.txt*).

Emplacement actuel des fichiers :

R:\ecophyto\Benjamin\Python\zon_stat\zonal_stats_fct_alti_pente_odr.py

R:\ecophyto\Benjamin\Python\zon_stat\fichier_de_config_odr.txt

Le module *Rasterstats* va dans un premier temps ajouter un champ « **count_mnt** » ; il s'agit du nombre de pixels de la *BD Alti* comptés dans chaque parcelle et pris en compte pour calculer les statistiques (un pixel est pris en compte dès lors que son centroïde intersecte la parcelle).

Champs de statistiques liés à l'altitude (exprimées en mètres) :

- « **a_min** » : altitude minimale de la parcelle
- « **a_max** » : altitude maximale de la parcelle
- « **a_mean** » : altitude moyenne de la parcelle
- « **a_std** » : déviation standard
- « **a_median** » : altitude médiane de la parcelle

Champs de statistiques liés à la pente (exprimées en degrés) :

- « **p_min** » : pente minimale de la parcelle
- « **p_max** » : pente maximale de la parcelle
- « **p_mean** » : pente moyenne de la parcelle
- « **p_std** » : déviation standard
- « **p_median** » : pente médiane de la parcelle

Champs de statistiques liés à l'exposition (Les valeurs sont comprises entre 0 et 360° et expriment la direction de la pente : à partir du nord (0°) et dans le sens horaire) :

- « **e_min** » : exposition minimale de la parcelle
- « **e_max** » : exposition maximale de la parcelle
- « **e_mean** » : exposition moyenne de la parcelle
- « **e_std** » : déviation standard
- « **e_median** » : exposition médiane de la parcelle

Pour plus de lisibilité, l'exposition est ensuite reclassée dans un nouveau champ « **exposition** » à partir du champ « **e_mean** » (Cf. *figure 7*) avec pour valeurs possibles :

- 'N' (0.0° ≤ « **e_mean** » < 11.25° ou 348.75° ≤ « **e_mean** » ≤ 360°)
- 'NNE' (11.25° ≤ « **e_mean** » < 33.75°)

- 'NE' (33.75° ≤ « e_mean » < 56.25°)
- 'ENE' (56.25° ≤ « e_mean » < 78.75°)
- 'E' (78.75° ≤ « e_mean » < 101.25°)
- 'ESE' (101.25° ≤ « e_mean » < 123.75°)
- 'SE' (123.75° ≤ « e_mean » < 146.25°)
- 'SSE' (146.25° ≤ « e_mean » < 168.75°)
- 'S' (168.75° ≤ « e_mean » < 191.25°)
- 'SSO' (191.25° ≤ « e_mean » < 213.75°)
- 'SO' (213.75° ≤ « e_mean » < 236.25°)
- 'OSO' (236.25° ≤ « e_mean » < 258.75°)
- 'O' (258.75° ≤ « e_mean » < 281.25°)
- 'ONO' (281.25° ≤ « e_mean » < 303.75°)
- 'NO' (303.75° ≤ « e_mean » < 326.25°)
- 'NNO' (326.25° ≤ « e_mean » < 348.75°)

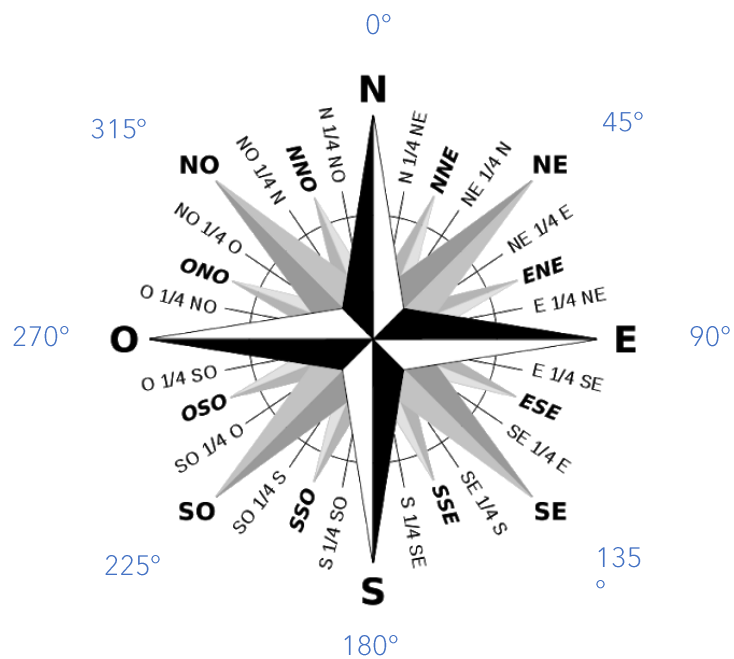


Figure 7 - Rose des vents – Correspondance orientation / angle

Note : les mêmes enrichissements concernant l'altitude, la pente et l'orientation des parcelles ont été réalisés sur les parcelles du RPG.

8. Conclusion

Comparé au *model builder* initial (Séard, 2016), les résultats sont améliorés :

- Les sources de données utilisées sont plus riches, actualisées et plus adéquates avec la temporalité du RPG à compléter.
- Les temps de traitement ont été divisés par 4 (1 mois => 1 semaine).
- Les *Parcelles ODR* sont maintenant enrichies de l'information sur les serres agricoles.
- Les *Parcelles ODR* sont également enrichies de l'information sur l'occupation du sol *OSO*.
- La prise en compte de l'intégralité du réseau routier permet d'éliminer certains axes qui étaient identifiés comme *Parcelles ODR* par le *model builder* (routes desservant les lotissements notamment).
- Concernant les 8 couches d'occupation du sol, la Version 3 de la *BD TOPO*® (largement enrichie par rapport aux versions précédentes) permet d'obtenir des résultats plus exhaustifs.
- La procédure est presque intégralement automatisée et reproductible, grâce à la portabilité et l'utilisation de logiciel libre (*Python*).

Limites et perspectives :

- La *BD Parcellaire* est encore partielle pour quelques départements (08, 11, 52, 55, 67 et 70). L'IGN travaille actuellement à compléter cette donnée.
- La généralisation des traitements à l'échelle nationale ainsi que la succession d'opérations spatiales complexes produisent ponctuellement des polygones aux géométries problématiques. Dans quelques rares cas, ces polygones bloquent les traitements malgré les « nettoyages » appliqués. L'erreur *Python* est alors enregistrée dans le fichier de log et le script passe au département suivant. Le programme reste cependant robuste puisque pour la campagne 2015, seulement 1 département sur 94 n'a pu être entièrement traité (1 pour 2016, 2 pour 2017 et 1 pour 2018).
- Il pourrait être envisagé de corriger automatiquement les géométries des polygones lorsqu'ils bloquent les traitements.
- Certaines *Parcelles ODR* comportent des « trous » qui compliquent les géométries. Ces trous sont issus de l'effacement d'arbres et de haies. Séard (2016) avait testé une méthode avec *QGIS* (non concluante) pour les effacer (tampon de 10 m suivi d'un tampon de -10 m).

9. Conclusion sur la phase 1 « *identification des parcelles agricoles hors RPG* » et perspectives

Les *parcelles ODR* peuvent être vues comme des parcelles cadastrales retravaillées (voir Annexe ci-dessous) ; l'utilisation des parcelles cadastrales et leurs découpages éventuels par des éléments artificialisés (chemins, routes, etc.) et des éléments paysagers (bois, forêt, haies, etc.) permet d'approcher correctement - d'après les observations de ce que nous avons produit - le concept de parcelles culturales et fournit des unités agricoles assez représentatives de la « réalité parcellaire » de l'occupation du sol agricole (hors RPG). Cet ensemble de parcelles identifiées est satisfaisant du point de vue des limites des « champs » ou des « prés ».

Il faut noter cependant la présence de parcelles agricoles trop découpées : la parcelle observée est représentée par plusieurs de nos parcelles, contiguës (*parcelle culturale est constituée de plusieurs parcelles cadastrales*). L'inverse est très rare (*la parcelle culturale est une subdivision d'une parcelle cadastrale*). Lors des observations et contrôles effectués, la fréquence d'apparition de cas où une de nos *parcelles ODR* englobe plusieurs parcelles culturales réelles et accueille plusieurs types de culture est inférieur à 1%.

Nous notons donc une adéquation satisfaisante entre la parcelle agricole (« odr ») et la parcelle cadastrale « retravaillée » (voir l'ensemble des figures présentées dans la partie 2 de la note

méthodologique), avec le bémol exposés ci-dessus (plusieurs de nos parcelles contigües forment une seule parcelle culturelle « réelle »). Ce fait revêt tout de même de l'importance lors de l'attribution des cultures (phase 2 de la méthode), car la superficie de la parcelle peut être corrélées avec la culture qu'elle accueille ; cette superficie est une des variables explicatives des modèles utilisés (voir à nouveau *Note méthodologique, partie 2*). Il y a donc un risque de prédire les cultures présentes sur plusieurs parcelles de « petites » tailles au lieu d'une seule culture sur une parcelle d'une taille « plus grande ».

La projection des fichiers géographiques est le RGF93/Lambert 93 (SCR EPSG : 2154) et sa résolution est celle des données IGN utilisée en entrée (Bd Parcellaire, Bd Topo®).

Ce fond parcellaire est à ce stade prêt pour la phase 2 de construction du RPG complété : attribuer une catégorie d'occupation du sol aux parcelles créées, pour écarter les dernières parcelles non agricoles et pour affecter une classe culturelle aux autres.

9. Références

Agreste, 2014. La Statistique Agricole Annuelle : présentation générale. Méthodologie, définitions et concept. Note méthodologique, Ministère de l'Agriculture et de l'Alimentation.
http://agreste.agriculture.sg-ppd.maaf.ate.info/IMG/pdf_methosaa.pdf

Cantelaube Pierre, Carles Marie, Séard Claire et Truche Camille, 2016. Méthodologie d'élaboration de la carte d'occupation du sol « RPG complété » version 1 (INRA ODR 2013-2017). Note méthodologique ODR, septembre 2016.

CCE Cour des Comptes Européenne, Rapport spécial, 2016, *Le système d'identification des parcelles agricoles: un instrument utile pour déterminer l'admissibilité des terres agricoles malgré une gestion encore perfectible*. Luxembourg: Office des publications de l'Union européenne, ISBN 978-92-872-5942-4, ISSN 1831-0850, doi:10.2865/416661.

CES "Occupation des SOIs" - OSO. <http://osr-cesbio.ups-tlse.fr/~oso/>

CLC, 2009, CORINE Land Cover France : Guide d'utilisation. Document technique du Commissariat général au développement durable Service de l'observation et des statistiques. Ministère de l'environnement, European Environment Agency. Février 2009

de Rincquesen Hubert, Julien Vincent , Lécuyer Charlotte, Levesque Robert, Liorit Dimitri, Melin Maxence, Pathier Guillaume et Quelin Claire, 2016, Structure foncière des exploitations agricoles. Appariement entre le registre parcellaire graphique et le cadastre pour construire une typologie des exploitations françaises d'après leur structure foncière. *Notes et Études Socio-Économiques (NESE) du Centre d'Etudes et de Prospective (CEP)*, n° 41, Décembre 2016, pp. 51-82.

IGN, 2018. RPG Version 1.0 – Descriptif de contenu et de livraison – Juillet 2018

IGN, 2019a. BD PARCELLAIRE® Version 1.2 – Descriptif de contenu – Juillet 2019

IGN, 2019b. Comparatif Parcellaire Express (PCI) / BD PARCELLAIRE® – Avril 2019

IGN, 2019c. BD CARTO® Version 3.3 – Descriptif de contenu – Juin 2019

IGN, 2020a. BD TOPO® Version 3.0 - Descriptif de contenu - Mai 2020

IGN, 2020b. RPG Version 2.0 – Descriptif de contenu et de livraison – Août 2020

INSEE, 2008, Code Officiel Géographique, www.insee.fr

INSEE, 2009, Les petites régions agricoles, www.insee.fr

ODR-Ecophyto, 2015. « *Mieux suivre spatialement l'usage des pesticides, en particulier sur les bassins versants, en s'appuyant sur un observatoire des ventes détaillé au code postal de l'utilisateur final de produit* » Rapport réalisé dans le cadre de l'axe 2, action 21 du programme Ecophyto, pour le compte de l'ONEMA et du MEDDE. Septembre 2015. INRA (US ODR, UMR Sadapt)

Séard Claire, 2016. Mise à jour de la cartographie d'occupation du sol, Note Méthodologique ODR, INRAE, septembre 2016.

Thines Georges et Lempereur Agnès, 1975, Dictionnaire général des sciences humaines, Paris, Éditions universitaires, 1034 pages.

10. Annexe - Notions de cadastre et de parcellaire agricole

Depuis le XIX^{ème} siècle, un plan parcellaire représente l'ensemble des parcelles d'une commune (le territoire de la commune est divisé en sections, chaque section donne lieu à un plan parcellaire). La parcelle est une portion de terrain appartenant à un même propriétaire et présentant une même nature (les plans peuvent faire apparaître des propriétés bâties - maisons et immeubles - et non bâties - prés, vignes, champs, etc.). Au cours du XX^{ème} siècle, des remaniements du cadastre sont intervenus dans le but de permettre des révisions des évaluations foncières servant de base à l'imposition. Aujourd'hui, les plans cadastraux sont régulièrement mis à jour, et sont consultables sous format numériques³. Réalisée à partir de l'assemblage de ces plans cadastraux dématérialisés, La BD Parcellaire® de l'IGN fournit l'information cadastrale numérique, géo-référencée et continue de l'ensemble du territoire français (IGN, 2019a). De première importance, cette base de données va constituer la donnée d'entrée sur laquelle repose notre couche d'information géographique, le « RPG complété ».

Cependant, la parcelle agricole (ou culturale) et la parcelle cadastrale présente des caractéristiques différentes : la 1^{ère} est définie via l'utilisation agricole dont elle fait l'objet alors que la 2nde est issue des informations relatives à sa propriété foncière.

En effet, la parcelle, selon Thines et Lempereur (1975), est *« une pièce de terre d'un seul tenant formant un tout au regard de la propriété, de l'exploitation ou de la culture; la parcelle est dite cadastrale lorsqu'elle est considérée sous l'angle juridique de la propriété, elle est qualifiée de culturale lorsqu'elle est envisagée sous l'angle de la culture qui la caractérise; elle est appelée parcelle d'exploitation, lorsqu'elle dépend d'un même exploitant. »*

Une information sur la nature des parcelles (terres agricoles, bois, constructibles, etc.) serait utile pour notre projet visant à compléter le RPG. Cette information repose sur une classification effectuée sur avis de la commission des impôts des communes et est disponible « uniquement » à ce jour sur demande aux mairies, d'après les numéros de parcelles. Elle n'est donc pas accessible pour une étude à l'échelle du territoire national comme la nôtre.

Les délimitations des parcelles culturales sont généralement moins stables dans le temps que celles des parcelles cadastrales. La parcelle culturale (un champ, un pré, une plantation, etc.) peut correspondre à un ensemble de parcelles cadastrales contiguës, entières ou partielles, portant généralement une seule culture, limitées par des éléments facilement repérables et permanents (tels qu'un chemin, un bois, un ruisseau, etc.). D'un autre côté, certaines parcelles culturales peuvent également découler d'une subdivision de la parcelle cadastrale. Alors qu'en est-il de nos « *parcelles ODR* » ? Sont-elles représentatives d'une réalité agricole sur le terrain ?

³ cadastre.gouv.fr



Pierre Cantelaube (pierre.cantelaube@inrae.fr)

INRAE, Centre Occitanie - Toulouse
Unité de Service de l'Observatoire du Développement Rural (0685)
24 chemin de Bordé Rouge, Auzeville - CS 52627
31326 Castanet-Tolosan Cedex, France

odr.inrae.fr

