

Constitution d'une base de données géographiques pour le suivi du parcellaire agricole à partir du registre parcellaire graphique

2 Mars 2020 – 19 Juin 2020

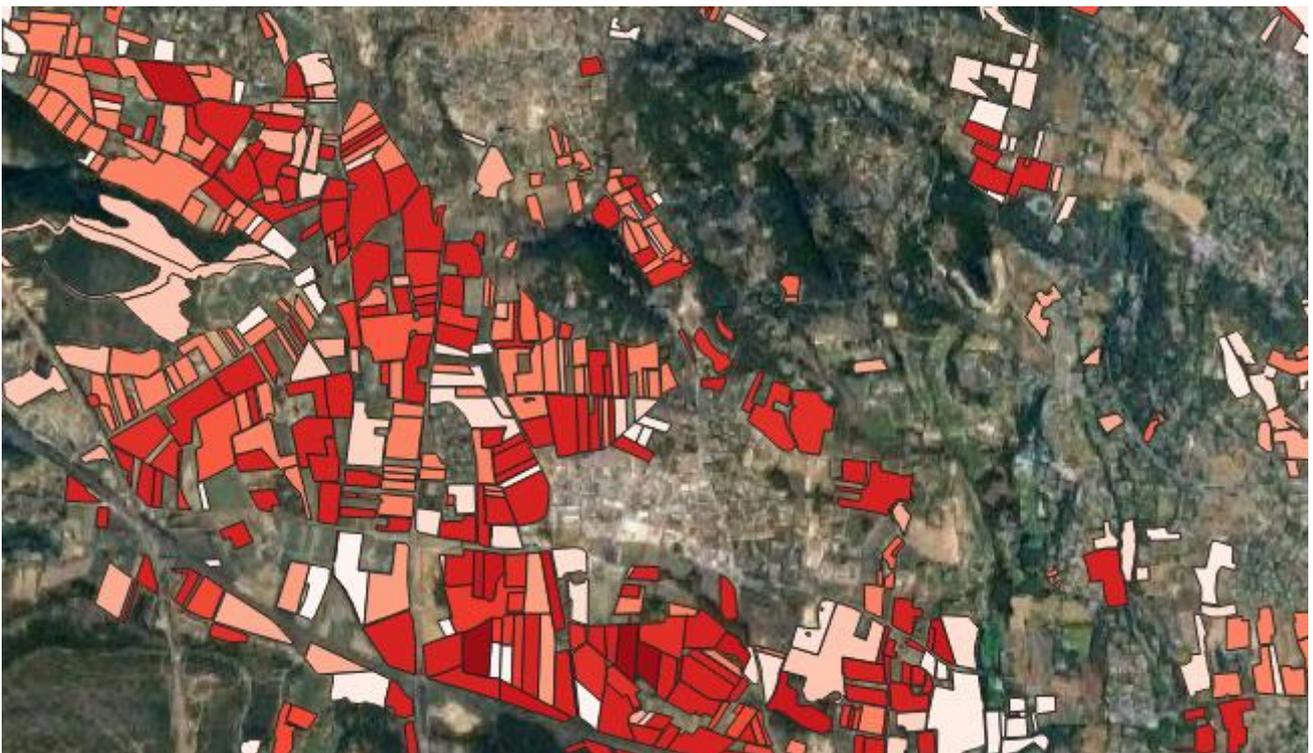


Figure 1: Extrait du Registre Parcellaire Graphique de 2015

Tuteur de stage : M. Pierre CANTELAUBE

Fanny HOUERIE

Enseignant référent : Mme Claire MARAIS-SICRE

Promo LP GGAT 2019-2020

Remerciements

Je tiens tout d'abord à remercier **Thomas Poméon**, directeur de l'ODR, pour m'avoir accueillie comme stagiaire au sein de son unité.

Je remercie vivement mon maître de stage, **Pierre Cantelaube**, pour m'avoir fait partager son expérience et son savoir, ainsi que pour sa disponibilité notamment dans ce contexte particulier du confinement.

Merci à **toute l'équipe de l'ODR**, pour sa disponibilité et sa bonne humeur qui ont rendu cette expérience enrichissante.

Table des matières

Introduction.....	5
I. Contexte et problématique	6
1. Présentation de la structure d'accueil.....	6
2. La plateforme ODR	7
3. Le Registre Parcellaire Graphique (RPG)	8
4. Projet	10
II. Moyens mis en œuvre / démarche	11
1. Le croisement du Registre Parcellaire Graphique avec les limites communales	11
2. Le croisement du Registre Parcellaire Graphique avec les sites Natura 2000	18
3. Optimisation des scripts	21
III. Résultats et discussion	23
1. Les tables de données	23
2. Les scripts	24
3. La note méthodologique	24
IV. Conclusion	25
1. Bilan pour l'entreprise.....	25
2. Bilan personnel.....	25
Bibliographie.....	26
Table des figures.....	27
Table des sigles.....	28
Annexes	29
1. Tableau de correspondance des groupes de cultures.....	29
2. Carte des régions Françaises et codes associés	30
3. Tableau des erreurs de géométries constatées sur le RPG	31
4. Scripts SQL	32
5. Résultats du croisement géographique de RPG avec les limites communales	36
6. Script SQL du croisement géographique RPG/ Communes (extrait).....	37
7. Script SQL du croisement géographique RPG/ Natura2000 (extrait).....	40
8. Note méthodologique	42

Introduction

L'unité de l'Observatoire du Développement Rural (ODR) propose l'accès, via sa plateforme web, aux données du Registre Parcellaire Graphique (RPG) pour les chercheurs et ingénieurs qui ne sont pas familiarisés avec les données géographiques et les outils SIG. Dans ce but, l'ODR prépare des tableaux où les données du RPG sont ramenées à des échelles administratives usuelles, la plus fine étant la commune Insee, pour une utilisation classique d'un tableau de données statistiques. Ces données correspondent essentiellement à des surfaces agricoles.

Les tableaux sur la plateforme ODR sont disponibles pour la période 2006-2014. A partir de la campagne 2015, le RPG a subi de nombreuses modifications (production par l'Institut National de l'Information Géographique et forestière, passage de l'îlot agricole à la parcelle agricole). De plus l'accès au RPG 2015 a été très tardif (disponible uniquement fin 2018).

L'objectif de mon stage était double :

- Mettre à jour la plateforme en intégrant les campagnes 2015 à 2017,
- Créer des scripts pour ajouter facilement les campagnes à venir.

Dans un premier temps, je vais vous présenter ma structure d'accueil, sa plateforme et mon outil de travail : le RPG. Dans un second temps, les étapes de réalisation de mon stage vous seront décrites. Et pour finir, je conclurais sur les résultats obtenus, le bilan pour l'entreprise et mon bilan personnel.

I. Contexte et problématique

1. Présentation de la structure d'accueil

INRAE (Institut National de la Recherche pour l'Agriculture, l'alimentation et l'environnement) est né le 1^{er} janvier 2020 de la fusion entre l'INRA (Institut National de la Recherche Agronomique) et l'IRSTEA (Institut National de Recherche en Science et Technologies pour l'Environnement et l'Agriculture). Cet institut a pour vocation d'accompagner l'évolution de l'agriculture en proposant de nouvelles orientations par la recherche, l'innovation et l'appui aux politiques publiques.

Le centre de recherche de Toulouse Midi-Pyrénées est un des centres de recherche de l'INRAE. Il est divisé en plusieurs unités dont l'Unité de Service de l'Observatoire du Développement Rural (US-ODR). (Figure 2)

Cette dernière est située sur le campus de Castanet-Tolosan (Sud-Est de Toulouse). Elle a été créée le 1^{er} janvier 2009 par l'INRA, le Ministère de l'Agriculture (MAAF) et l'Agence de Services et Paiements (ASP). Sa mission est d'assurer le développement et la gestion de la plateforme de « l'Observatoire des programmes communautaires de Développement Rural » (ODR).

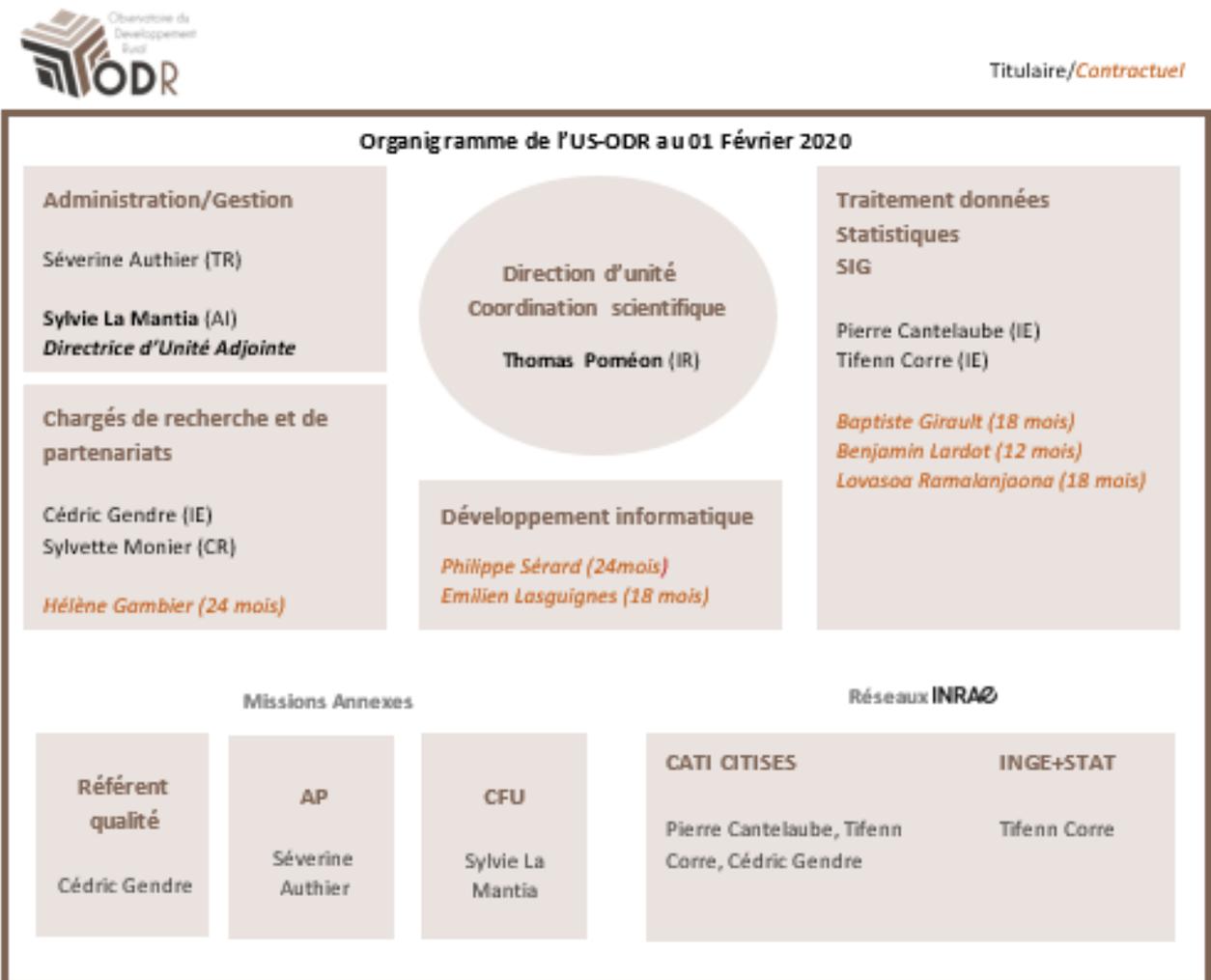


Figure 2: Organigramme de l'US-ODR Source: ODR

3. Le Registre Parcellaire Graphique (RPG)

Le Registre Parcellaire Graphique (RPG) est un système d'informations géographiques permettant l'identification des parcelles agricoles sur le Territoire français. Il repose sur un système de déclarations graphiques effectuées par les agriculteurs dans le cadre de la Politique Agricole Commune (PAC).

Une convention a été établie entre l'INRA et l'ASP pour la mise à disposition des données du RPG de 2006 à

2014, dans une version anonyme (niveau 4), pour les travaux de recherche de l'INRAE. Aujourd'hui, c'est l'IGN (Institut National de l'Information Géographique et forestière) qui met à disposition la version 2.0 du RPG (Figure 4). La convention initiale est donc caduque. Les données utilisées par l'US-ODR sont alors les données publiques de niveau 2.

Les données obtenues à partir de 2015 sont proposées dans les systèmes de référence ci-contre. (Figure 5)

L'information disponible est donc celle des parcelles des exploitations qui font l'objet d'une aide de la PAC. Jusqu'en 2014, les données étaient fournies tous les ans, sous formes géographiques vectorielles (format ESRI Shapefile), avec des données attributaires associées : la table des informations relatives à la culture (nature, surface...) et celle concernant l'exploitation (identifiant unique anonyme, structure de l'exploitation surface totale, communes...). A partir de 2015, l'IGN fournit les données dans un unique fichier sous formes géographiques vectorielles (format Shapefile ESRI). Les informations concernant l'exploitation et la commune ne sont plus présentes dans les attributs (Figure 6) et l'échelle initiale n'est plus l'îlot agricole mais la parcelle agricole.

Niveau 1	Données graphiques avec un identifiant numérique et non significatif (anonyme) par îlot
Niveau 2	Données du niveau 1 avec la commune de localisation de l'îlot, les cultures déclarées et leurs surfaces décrites en 28 groupes de cultures (occupation du sol)
Niveau 3	Données du niveau 2 avec pour chaque îlot, sa surface de référence, s'il y a un moins une parcelle irriguée ou non, et les caractéristiques de l'exploitation (anonymisées) : forme juridique, classe d'âge pour les exploitants individuels, surface déclarée, département de rattachement administratif
Niveau 4	Données du niveau 3 avec par exploitation un identifiant numérique et non significatif
Niveau 5	Données du niveau 4 avec nom ou raison sociale de l'exploitant
Niveau 6	Données du niveau 5 avec régime et montant des aides

Figure 4: Tableau des différents niveaux d'anonymisation du RPG. Source: ODR

Zone	Système géodésique	Ellipsoïde associé	Projection	Unité	Résolution
France métropolitaine	RGF93	IAG GRS 1980	Lambert 93	m	cm
Guadeloupe, Saint-Barthélemy, Saint-Martin, Martinique	WGS84	IAG GRS 1980	UTM Nord fuseau 20	m	cm
Guyane	RGFG95	IAG GRS 1980	UTM Nord fuseau 22	m	cm
La Réunion	RGR92	IAG GRS 1980	UTM Sud fuseau 40	m	cm
Mayotte	RGM04	IAG GRS 1980	UTM Sud fuseau 38	m	cm

Figure 5: Tableau des projections du RPG. Source: IGN

Attributs	Définition	Caractéristiques
<i>Id_parcel</i>	Identifiant unique de la parcelle	Caractères (max 10)
<i>Code_cultu</i>	Code de la culture principale	Caractères (3), présent en annexe
<i>Code_group</i>	Code du groupe de la culture principale	Caractères (2), présent en annexe
<i>Culture_d1</i>	Code de la culture dérobée 1	Caractères (3), vide si nul
<i>Culture_d2</i>	Code de la culture dérobée 2	Caractères (3), vide si nul
<i>Surf_parcel</i>	Surface en hectare de la parcelle	Décimal (8,2)

Figure 6: Données attributaires du RPG version 2.0

L'identifiant unique de la parcelle est conservé d'une année sur l'autre pour les années postérieures à 2015, lorsque la géométrie de la parcelle n'est pas modifiée.

Les cultures sont définies par un code (« code_cultu ») et sont rassemblées par groupes qui portent chacun un numéro (« code_group »). (Annexe 1)

La distribution de ces données « brutes » (Version 0) aux équipes INRAE est la première mission du Réseau Systèmes Agricoles (RSA). La seconde mission est d'enrichir ces données (versions 1 et 2), grâce à des ressources issues directement ou indirectement du RPG, afin d'améliorer les travaux d'évaluation et de modélisation des systèmes agricoles.

Les versions 1 et 2 correspondent respectivement aux versions « RPG enrichi » et « RPG complété » qui peuvent être fournies par l'observatoire. Le RPG « enrichi » comporte, en plus du RPG « brut », les données des Mesures Agro-Environnementales (MAE). Cela permet d'utiliser le RPG dans des études liées aux problématiques environnementales dans l'agriculture. Le RPG « complété » essaie de répondre à la problématique des terrains agricoles non répertoriés dans le RPG car non éligibles aux aides de la PAC. Pour cela, les zones potentielles agricoles sont détectées à partir de la BD Topo (IGN) et validées ou non par des outils statistiques.

Des indicateurs tels que celui des systèmes de cultures (nombre de culture, répartition dans la surface agricole utile (SAU)...) sont modélisés dans la plateforme. Il est possible de faire varier les paramètres d'échelles (commune, petite région agricole, département et région), les types d'occupations du sol... et d'exporter les cartes ou le résultat au format « .xsl » lorsqu'il s'agit de tableaux. Des indices de diversité des cultures sont aussi présents : indices de Shannon et de Simpson.

La base de données qui alimente la partie « RPG » de la plateforme de l'ODR est non-géographique : MySQL. Elle a pour but de créer des tableaux statistiques pour les personnes non familiarisées avec les outils SIG (systèmes d'Informations Géographiques). Lors de la création de ces tableaux, plusieurs paramètres peuvent être modifiés. Le premier est l'échelle géographique : du référentiel communal à la France Métropolitaine, en passant par les régions agricoles, les départements, les régions... Le second

Introduction | **Notes de mise à jour** | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013

Surfaces RPG

Etendue géographique: **France entière**

Modifier le niveau d'agrégation des indicateurs de ce tableau

Le niveau actuel d'agrégation des indicateurs est **Départements**
 Vous pouvez cependant modifier ce niveau d'agrégation en choisissant un parmi la liste suivante:

★ Vous possédez un droit de consultation qui vous permet de consulter ces indicateurs jusqu'au niveau : **Communes**

📘 Pour plus d'informations sur les niveaux géographiques disponibles, consultez le [catalogue des données géographiques](#)

📩 Si vous souhaitez formuler une demande pour avoir accès à un niveau plus fin, vous pouvez le faire [en cliquant ici](#)

Paramètres du filtre pour les indicateurs

Types d'indicateurs à afficher:

Types de cultures à afficher:

Filtre sur les exploitations:

Contenu de ce tableau d'indicateurs

Indicateurs de surfaces RPG (en milliers d'ha) par Départements
 Indicateurs de surfaces RPG en taux par Départements

Indicateurs de surfaces RPG (en milliers d'ha) par Départements

Note :

- Dans le tableau ci-dessous apparaissent les 28 groupes de cultures du regroupement initial ASP.
- Pour plus de détails, merci de consulter la note méthodologique de ce dossier.

Chercher

code	Départements	Blé tendre	Maïs grain et ensilage	Orge	Autres céréales	Colza	Tournesol	Autres oleagineux	Proteagineux	PI à
01	Ain	29.96	68.54	7.25	6.63	6.49	4.49	3.44	0.36	
02	Aisne	189.45	35.02	41.68	3.64	52.63	0.30	0.15	12.16	

Figure 7: Vue de la plateforme web ODR

concerne les indicateurs : les indicateurs de surfaces, les indicateurs de surface avec contexte Natura2000 et les indicateurs de surface et de diversité. (Figure 7)

Pour alimenter la base de données il est alors nécessaire de rapporter toutes les tables à l'échelle initiale : la commune.

4. Projet

L'objectif de l'US-ODR est de permettre aux personnes qui ne sont pas familières avec les données géographiques et les outils SIG d'accéder aux données du Registre Parcellaire Graphique (RPG)

L'ODR prépare des tableaux où les données du RPG sont ramenées à des échelles administratives usuelles, la plus fine étant la commune Insee, pour une utilisation classique d'un tableau de données statistiques. Ces tableaux sont interactifs et l'utilisateur de la plateforme peut agréger les données (c'est-à-dire des surfaces agricoles par type de culture) à des échelles compatibles avec le référentiel communal (cantons, départements, etc.) ou encore délimiter une zone d'intérêt particulière (Régions, bassins versants, territoires agricoles, etc. ; critères géographiques qui doivent néanmoins être compatibles avec le référentiel communal).

Ces tableaux sur la plateforme ODR sont disponibles pour la période 2006-2014.

L'objectif de mon stage est de mettre à jour la plateforme web de l'ODR en proposant les campagnes 2015 à 2017, ainsi que de créer des scripts pour ajouter facilement (et le plus automatiquement possible) les campagnes à venir. Pour cela, il faut nécessairement associer à chaque parcelle du RPG la commune dans laquelle elle se trouve, afin de préparer les tables de données qui alimenteront les tableaux décrits ci-dessus.

D'autre part, la connaissance des surfaces RPG présentes dans les sites du Réseau Natura2000 est particulièrement intéressante (agroenvironnement, transition agro-écologique, ...). L'ODR propose ainsi de recenser ces surfaces dans ces tableaux de données relatifs au RPG (du niveau parcellaire aux échelles administratives de la plateforme). Il est donc nécessaire de croiser géographiquement les parcelles du RPG et les limites des sites Natura2000.

En résumé, le RPG est un outil cartographique essentiel pour disposer, sur l'ensemble du territoire national, d'une base de données qui permet de croiser des informations sur les caractéristiques des parcelles (surfaces, localisation) et des pratiques culturelles.

Ces données seront indispensables pour toutes les recherches et analyses nécessitant une connaissance fine de l'évolution des pratiques agricoles et de l'occupation de l'espace par des activités agricoles (SAU, types de cultures...).

On peut aisément imaginer la diversité et la multiplicité des personnes et organismes, publics ou privés, susceptibles de bénéficier d'un accès à une telle base de données.

L'accès à un public aussi diversifié, y compris à des non professionnels (étudiants, journalistes, collectivités locales...), va nécessiter de vulgariser l'information notamment pour les personnes qui ne manient pas les SIG.

II. Moyens mis en œuvre / démarche

L'objectif qui m'avait été donné consistait à actualiser la plateforme de l'ODR avec les données du RPG de 2015 à 2017.

Pour cela, le RPG devait être croisé géographiquement avec différentes couches de données géographiques telles que les limites communales et la couche Natura2000. Ces croisements seront effectués sur le logiciel PostgreSQL avec l'extension PostGis en utilisant le langage SQL.

J'ai effectué l'ensemble des traitements présentés ci-dessous à trois reprises, une fois pour chaque année.

1. Le croisement du Registre Parcellaire Graphique avec les limites communales

Dans la base de données de l'ODR, le RPG est présent en trois groupes de fichiers principaux : un pour chaque année (2015 à 2017) ; qui contiennent chacun cinq fichiers : un pour chacune des projections (Guadeloupe/ Martinique, Guyane, Réunion, Mayotte et Métropole), soit un total de 18 fichiers.

La couche métropole existe aussi en 13 couches distinctes qui correspondent aux 13 nouvelles régions (annexe 2).

Nous travaillerons uniquement sur les couches « régions » de la métropole car la plateforme de l'ODR n'est pas programmée pour le traitement des DOM-TOM.

Chaque année, ce sont environ 9 400 000 parcelles qui sont déclarées, en métropole, dans le cadre de la PAC. Une parcelle est définie par un identifiant unique (« id_parcel »), sa surface (« surf_parc »), son code culture (« code_cultu »), le code du groupe (« code_group ») et sa géométrie (« wkb_geometry »). De plus, lors de l'import du fichier au format « Shape » sur PostgreSQL, un identifiant unique est ajouté (« ogc_fid »). Il correspond à la clé primaire. Deux champs sont aussi prévus au cas où la parcelle accueillerait également des cultures dérobées (culture_d1 et culture_d2). Une culture dérobée se place entre deux cultures principales, entre deux saisons.

Il faut donc commencer par attribuer à chaque parcelle le code Insee de la commune où elle est située. Pour cela il est nécessaire d'intersecter les couches RPG avec la couche des communes.

Lors de la superposition de ces deux couches, nous avons pu remarquer que de nombreuses parcelles étaient situées sur la limite de deux, voire plusieurs communes. Nous pouvons donc émettre deux hypothèses : soit le fichier des limites communales à une résolution moins fine que celle du RPG soit les parcelles sont réellement « à cheval » sur deux communes. (Figure 8)

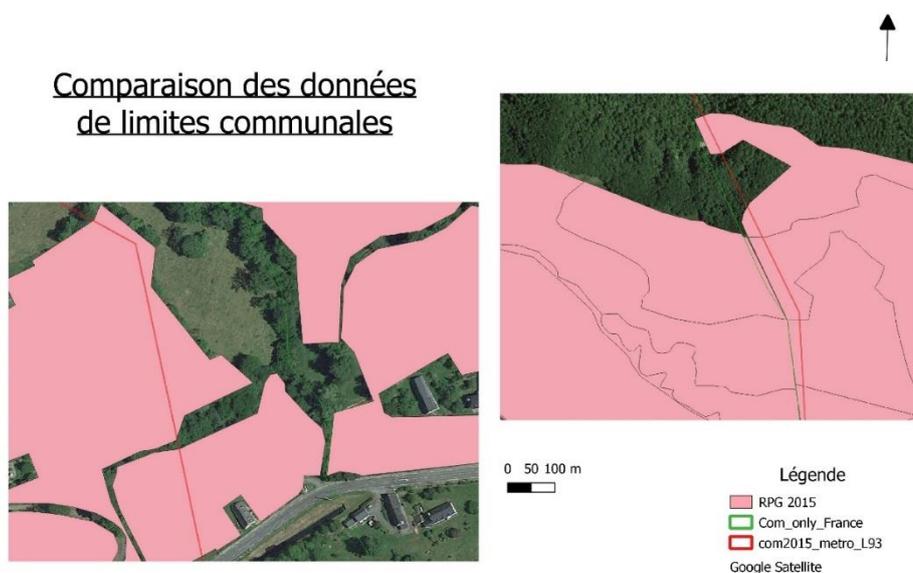


Figure 8: Carte de comparaison de la résolution des deux sources de limites communales, ainsi que des limites des parcelles RPG.

Afin d'évaluer la première hypothèse, nous avons choisi d'utiliser un autre fichier des limites communales d'une résolution plus fine, présent dans la base de données de l'ODR (« *com_only_france* », construit par agrégation des limites communales de la Bd Topo® de l'IGN).

On observe, comme dans le cas ci-contre (Figure 9), une nette amélioration même si quelques erreurs sont encore visibles. Après réflexion nous avons choisi d'utiliser la couche « *tb_commune_2014_metropole* » obtenue à partir de la fusion des polygones de la couche IRIS (Ilots Regroupés pour l'Information Statistique). Cette couche a été développée par l'INSEE en 1999 pour les besoins de recensement de la population. Elle correspond à un découpage du territoire en maille de taille homogène. Cette couche est plus précise qu'une couche communale puisqu'elle est à une échelle inférieure (base à l'échelle du 1 : 1 000 000).

La couche IRIS utilisée pour les croisements avec les RPG de 2015, 2016 et 2017 est celle de 2014. Elle a été réalisée avec les résultats du recensement de la population 2011. (Source INSEE)

Malgré une plus importante précision, on peut voir que la seconde hypothèse est aussi vérifiée : certaines parcelles sont sur plusieurs communes.

Nous allons donc attribuer à chaque parcelle le code Insee de la commune sur laquelle la surface de la parcelle est majoritaire.

L'import des couches nécessaires aux traitements sur le logiciel PostgreSQL

Après de nombreux essais avec les couches du RPG à l'échelle de la métropole, nous avons constaté qu'elles étaient trop lourdes pour effectuer de tels traitements. Nous avons donc utilisé les couches du RPG déjà divisées par région. Ainsi, il sera nécessaire de réaliser 13 fois le même traitement : une fois par région (annexe 2).

Pour importer ces fichiers sur PostgreSQL il existe deux solutions :

- Soit manuellement via PostgreSQL ou l'interface « base de données » de QGIS
- Soit via l'invite de commande QGIS (OSGeo 4W Shell)

La première solution étant très lente, c'est la seconde qui a été choisie. Une ligne de code a été écrite afin d'extraire les fichiers de type Shapefile d'un dossier et de les ajouter à ma base de données sur PostgreSQL :

```
ogr2ogr -f "PostgreSQL" PG:"host=localhost user=postgres dbname=rpg password=ggat" -nlt MULTIPOLYGON
-a_srs EPSG:2154 -s_srs EPSG:2154 E:/rpg2016/RPG_2-0_SHP_LAMB93_R11-2016/PARCELLES_GRAPHIQUES.shp
-nln rpg2016.tb_rpg_2016_reg_11_ile_de_france
```

« Ogr2ogr » est la fonction qui permet de convertir le format d'un fichier. Plusieurs options peuvent être ajoutées :

Carte représentant le RPG 2015 et différentes tables de données des limites communales

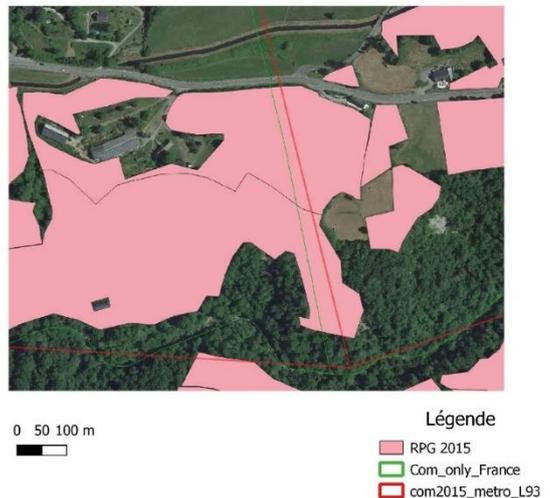


Figure 9: Carte de comparaison des limites communales sur le RPG de 2015.

- « -f » : correspond au format de sortie du fichier, dans notre cas PostgreSQL.
- « PG » permet d'accéder à la base de données.
- « -nlt » : définit le type de géométrie. Le RPG est composé de parcelles qui peuvent être multi-parties. Le type de géométrie est donc multi-polygone.
- « -s_srs » et « -a_srs » : permettent respectivement d'écraser la projection et de la redéfinir. Dans notre cas, la projection originale était correcte mais elle n'était pas reconnue.
- Ensuite il était nécessaire d'indiquer le chemin du fichier à utiliser.
- « -ln » permet d'assigner un nom à la nouvelle couche.

C'est ainsi que les 13 couches par année sont exportées, de même pour la couche des communes. Une clé primaire est ajoutée automatiquement sur un nouveau champ « ogc_fid » et la colonne géométrie est indexée.

Correction des géométries

Les données du RPG sont des couches vectorielles composées de multi-polygones dessinés par des personnes non-issus des métiers de la géomatique, principalement des agriculteurs. Il est donc possible que, lors du traçage des polygones, des erreurs soient faites, principalement des auto-intersections et des doublons.

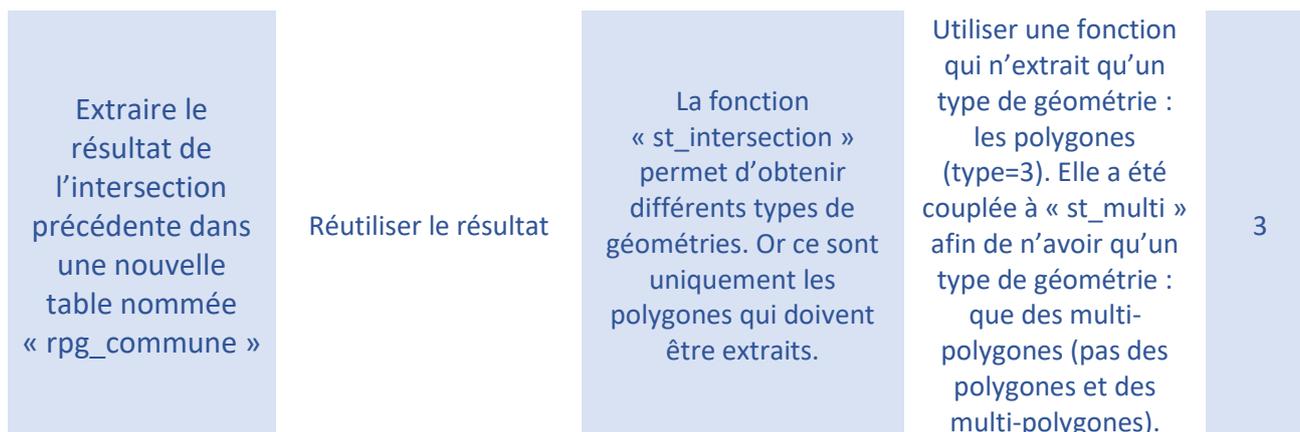
Pour détecter ces erreurs, c'est la fonction « ST_IsValid » qui est utilisée. Par exemple, pour l'année 2015, 53 erreurs étaient présentes au total sur les régions (annexe 3). Si des erreurs sont détectées elles doivent être réparées et c'est la fonction « ST_MakeValid » qui le permettra. Celle-ci crée une représentation valide d'une géométrie non valide donnée, sans perdre aucun des sommets d'entrée. Les géométries déjà valides sont retournées sans aucune intervention. Les polygones simples peuvent devenir des multi-géométries en cas d'auto-intersection. Nous allons donc créer une table « rpg_region » qui contient l'ensemble des informations de la table du RPG de départ avec les géométries modifiées (annexe 4, script 1)

Il est important de redéfinir la clé primaire sur le champ « ogc_fid » et le SRID (Spatial Reference Identifier) qui est dans notre cas 2154 (Lambert 93). Ensuite, nous pouvons indexer une ou plusieurs colonnes afin d'optimiser les prochains traitements.

Associer à chaque parcelle le code Insee de la commune sur laquelle elle se situe

Pour cela il faut intersecter les couches du RPG avec celles des limites communales afin de récupérer la géométrie de l'intersection et les identifiants « depcom » (code INSEE de la commune) ainsi que l'identifiant unique de la parcelle (« id_parcel »). La table créée lors de cette intersection (« rpg_commune ») pourra ainsi facilement être jointée aux tables de base afin de récupérer les autres informations. (Annexe 4, scripts 2 et 3)

<i>Etapes</i>	<i>Objectif</i>	<i>Problèmes rencontrés</i>	<i>Solutions</i>	<i>Script</i>
Intersecter le RPG et la couche des limites communales	Découper le RPG selon les limites communales et affecter à chaque partie du RPG un code INSEE	Si les géométries ne partagent aucun espace, une collection vide est renvoyée.	Effectuer l'intersection que si les géométries A et B sont jointes.	2



Traitements sur la couche « rpg_commune »

Calcul de la surface des parcelles intersectées

Le champ « surf_parce » correspond à la surface initiale de la parcelle exprimée en hectares avec une précision de 0.01 ha, donnée par l'IGN. Il faut donc calculer la nouvelle surface des polygones avec la fonction « ST_Area ». Le résultat de cette fonction est en m², il faut donc diviser par 10 000 pour obtenir la surface en hectares.

La surface est stockée dans un champ nommée « surface » de type numérique, avec 5 décimales soit une précision de 0.1m². (Annexe 4, script 4)

Attribuer un code Insee unique à chaque parcelle

Dans un premier temps, il avait été choisi de calculer le pourcentage de surface de la parcelle découpée par rapport à la surface initiale de la parcelle pour savoir si l'intersection avait découpé la parcelle apparaissant sur deux communes ou si la parcelle a été découpée à cause d'erreurs de digitalisation des limites communales. Pour cela, nous avons créé une nouvelle colonne dans laquelle nous voulions calculer ce pourcentage :

```
UPDATE fanny.tb_rpg_2015_com_bourgogne_franche_comte
set pourc_surf_tot = ((surface/surf_parce) * 100)
```

Or, nous avons été confrontés à plusieurs problèmes.

Le premier était que les très petites surfaces de 0 ha (à cause de l'arrondie à 2 chiffres après la virgule de la surface en hectare) ne permettaient pas de calculer un pourcentage. Pour résoudre cette difficulté nous avons augmenté la précision de la colonne surface en ajoutant une nouvelle colonne (« surface_ge », présentée ci-après).

Le second problème consistait à trouver une solution pour les parcelles divisées sur plus de deux communes.

Le troisième nous obligé à choisir un pourcentage à la limite entre celui lié à une erreur de digitalisation et celui lié à une réelle limite communale. Nous n'avons pas pu établir ce pourcentage « limite ». En effet, en fonction de la surface totale de la parcelle, celui-ci pouvait varier du simple au double. Nous avons donc choisi

d'attribuer à la parcelle le code Insee de la commune sur laquelle l'emprise était majoritaire. C'est cette technique qui avait été appliquée par l'ASP jusqu'en 2014.

<i>Etape</i>	<i>Objectif</i>	<i>Script</i>
Ajouter une colonne « surf_maxi » qui contiendra la surface du « morceau de parcelle » qui a la surface maximale	Trouver sur quelle commune la parcelle est majoritairement	5

Dans le cas où la parcelle est sur une seule commune alors la surface sera égale à la surface maximale. Il suffira alors de faire une jointure avec les lignes de la table où la surface de l'intersection avec les communes (« surface ») est égale à cette « surf_maxi ». Pour finir, la table temporaire est supprimée ce qui permet d'éviter de surcharger le serveur (annexe 4, script 5).

Traitements sur la couche « rpg_region »

<i>Etapes</i>	<i>Objectif</i>	<i>Problèmes rencontrés</i>	<i>Solution</i>	<i>Script</i>
Ajouter une colonne « nbr_com » qui compte le nombre de commune que la parcelle intersecte	Savoir si la parcelle est sur une ou plusieurs communes			6
Ajouter une colonne « code_insee » et récupérer le code Insee de la commune où le plus grand « morceau de parcelle » est situé, c'est-à-dire où « surface » = « surf_maxi »	Attribuer à chaque parcelle initiale un code Insee	Lorsqu'il existe deux surfaces maximales égales il ne sera pas possible de récupérer le code Insee. A l'inverse lorsque la parcelle n'est pas superposée avec la couche des limites communales alors elle n'a aucun code Insee.	Expliquée ci-après	Ci-dessous

```
ALTER TABLE rpg2016.tb_rpg2016_reg_l1_ile_de_france ADD COLUMN code_insee character varying (6);
update rpg2016.tb_rpg2016_reg_l1_ile_de_france set code_insee =
(select a.code_insee
 from rpg2016.tb_rpg_2016_com_l1_ile_de_france as a
 where a.id_parcel = rpg2016.tb_rpg2016_reg_l1_ile_de_france.id_parcel and a.surface= a surf_maxi );
```

Les deux cas particuliers relevés lors de la précédente étape ont été traités séparément.

Il arrive qu'une parcelle soit divisée en deux surfaces égales à 0.1m² prêt. Dans ce cas, le script de l'étape précédente (ci-dessus) renvoie une erreur libellée « plus d'une ligne renvoyée par sous requête ». Ce qui signifie que plusieurs codes Insee sont renvoyés pour une unique parcelle. Pour résoudre ce problème, nous avons choisi d'attribuer aux parcelles concernées le code Insee le plus petit, qui correspond à la première des deux communes par ordre alphanumérique. Pour cela, une colonne qui compte le nombre de surface égale a été ajoutée (« nb_surf_mm »). Dans le cas où « nb_surf_mm » est égal à 1 alors c'est le processus précédent

qui est appliqué. Et dans le cas inverse : si « nb_surf_mm » est supérieur à 1, alors le script attribue le code Insee minimum. (Script ci-dessous)

```
---ajout d une colonne nb_surf_mm qui compte le nombre de morceau de parcelle qui ont la meme surface
Alter table schema.tablecom add column nb_surf_mm numeric (5,1);
update schema.tablecom set nb_surf_mm =
  (select count (distinct a.code_insee)
   from schema.tablecom as a
   where a.surface= a.surf_maxi and a.id_parcel=schema.tablecom.id_parcel);

---completer la colonne code_insee pour les cas normaux: nb_surf_mm = 1 ---
ALTER TABLE schema.tablereg ADD COLUMN code_insee character varying (6);
update schema.tablereg set code_insee =
  (select a.code_insee
   from schema.tablecom as a
   where a.nb_surf_mm=1 and a.id_parcel = schema.tablereg.id_parcel and a.surface= a.surf_maxi );

---completer la colonne code_insee pour les cas où nb_surf_mm > 1 ---
update schema.tablereg set code_insee = (
  WITH tb3 as (
    SELECT tbl.id_parcel
    FROM schema.tablereg as tbl),
    tb4 as (
      select min(code_insee) as codinsee, a.id_parcel as parcel, a.nb_surf_mm
      from schema.tablecom as a
      where nb_surf_mm>1
      group by a.id_parcel, a.nb_surf_mm
    )
  SELECT tb4.codinsee
  FROM tb3
  LEFT JOIN tb4 ON tb3.id_parcel=tb4.parcel
  where tb4.nb_surf_mm>1 and tb3.id_parcel=schema.tablereg.id_parcel
)where code_insee is null;
```

Ce cas est rencontré en moyenne 5 à 6 fois sur 2015, 2016 et 2017, dans trois régions différentes. (Annexe 5)

Le second cas concerne les parcelles qui n'étaient pas dans la zone définie par la couche des communes. Dans ce cas, la parcelle n'est pas supprimée mais elle ne possède pas de code Insee. Ces cas rares concernent principalement des parcelles en Espagne, Allemagne, Belgique ou Luxembourg qui sont déclarées par des exploitants agricoles « frontaliers » ou encore des regroupements pastoraux entre français et espagnol pour les estives. Certaines parcelles qui sont en bordure de côte ou sur des îles non-répertoriées par la couche des communes sont aussi perdues. Ce problème concerne entre 550 et 600 parcelles par an soit 0.006% ce qui est minime à l'échelle du RPG. Il a donc été décidé de ne pas garder ces parcelles car compléter manuellement le code Insee de celles-ci aurait été très chronophage.

Enfin, nous avons choisi de recalculer la surface en hectare des parcelles initiales du RPG. Cette information est déjà présente dans les tables transmises par l'IGN mais il se peut qu'elle soit erronée. De plus la précision sera au m² alors qu'elle était de 0.01 ha dans la source. Le nom de cette colonne est « surface_ge », ce qui signifie « surface géographique ».

Pour agréments la base de données de l'ODR il est nécessaire d'unir toutes les tables précédentes afin d'obtenir une unique table pour la France entière.

Or, les parcelles qui chevauchent la limite de deux régions sont présentes dans les deux tables. Comme dans le cas présenté ci-contre (Figure 10), des parcelles sont à la fois en rose et en violet ce qui signifie qu'elles sont dans la table du RPG de l'Occitanie et dans celle de Nouvelle Aquitaine.

Lors de l'union de toutes les tables ces parcelles sont comptées autant de fois qu'elles sont sur une région. Nous avons donc choisi de conserver chaque parcelle sur la région où sa surface est majoritaire.

Pour cela, une nouvelle table a été créée avec les parcelles dont le code Insee appartient à un des départements de la région. (Annexe 4, script 7)

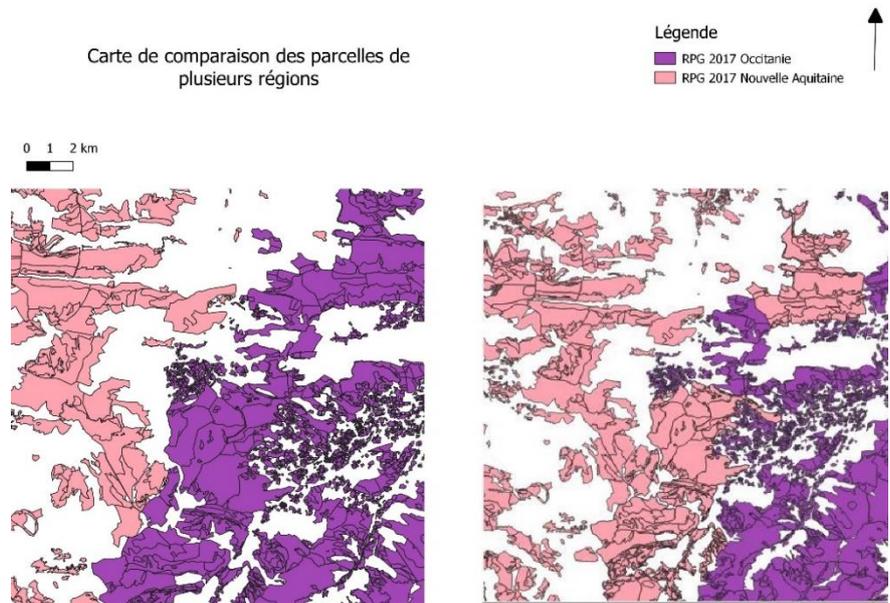


Figure 10: Carte de comparaison des parcelles du RPG 2017 (à gauche : Occitanie au 1^{er} plan et Nouvelle-Aquitaine au 2nd plan ; à droite l'inverse)

Cette erreur de superposition provoquait une augmentation de surface de 150 000 ha pour le RPG 2015.

Pour finir, toutes les tables ont été réunies grâce à la fonction « union all » afin d'obtenir, pour chaque année, une ligne par code de groupe de culture et par commune avec la surface représentée par ce « code_group ». (Annexe 4, script 8)

Le résultat est le suivant : (Figure 11)

	code_insee character varying (6)	code_group character varying (2)	surf_ha numeric
1	2A001	16	28.74
2	2A001	17	256.96
3	2A001	18	248.82
4	2A001	19	13.91
5	2A001	20	1.55
6	2A001	21	2.62
7	2A001	25	0.40
8	2A001	28	11.73
9	2A004	11	4.33
10	2A004	16	18.28

Figure 11: Vue d'une partie de la dernière table fournie

Cette table est donc non-géographique (c'est un tableau de données classique, avec néanmoins une référence géographique : la commune Insee). Elle est exportée au format « .csv » afin d'être traitée sur le logiciel PhpMyAdmin (Système de gestion de base de données MySQL, support de la plateforme web ODR).

Le traitement sur ce logiciel a été effectué par mon maître de stage. Il vise à transformer la table afin d'obtenir une ligne par commune. C'est alors cette table qui a été importée sur la base de données de l'ODR afin d'agrémenter la fiche du réseau « système agricole » de la plateforme.

2. Le croisement du Registre Parcellaire Graphique avec les sites Natura 2000

Afin d'agrémenter la base de données de l'ODR, les données du RPG ont été croisées avec celles de la couche des limites géographiques des sites Natura 2000 (www.natura2000.fr).

Le réseau européen *Natura 2000* regroupe les outils de la politique européenne de préservation de la biodiversité. Il est constitué d'un ensemble de sites naturels, terrestres et marins qui ont pour objectif de préserver la diversité biologique et le patrimoine naturel.

Natura 2000 en 2019, quelques chiffres :

- Natura 2000 en Europe :
 - 18.15 % de la surface terrestre du territoire de l'Union européenne ;
 - 6 % de la surface marine des eaux européennes ;
 - 5 572 zones de protection spéciale pour les oiseaux (ZPS) ;
 - 23 726 zones spéciales de conservation pour les habitats et les espèces (ZSC ou SIC) ;
- Natura 2000 en France :
 - 12.9 % de la surface terrestre métropolitaine, soit 7 millions d'hectares ;
 - 33 % de la surface marine de la zone économique exclusive, soit 12 millions d'hectares ;
 - 1 776 sites, dont 212 sites marins : 402 zones de protection spéciales pour les oiseaux (ZPS) et 1 374 zones spéciales de conservation (ZSC ou SIC)

Dans notre cas, ce sont les zones terrestres qui vont être intersectées avec le RPG. D'après le site du ministère de la transition écologique et solidaire, le réseau terrestre se répartit sur les milieux suivants :

- 43 % de forêts ;
- 29 % de prairies et landes ;
- 20 % de zones agricoles cultivées ;
- 4 % d'habitats rocheux (roches nues, plages, dunes...) ;
- 3 % de cours d'eau, tourbières et marais ;
- 1 % de zones urbaines.

Les prairies et landes devraient correspondre principalement aux codes cultures des prairies permanentes (PPH et PRL dans le RPG) et des estives et landes (BOP, SPH et SPL). Quant aux zones agricoles cultivées, elles peuvent se situer sur toutes les autres catégories de cultures.

La couche du RPG 2015 est croisée avec la version de septembre 2014 des sites Natura 2000. Alors que les RPG 2016 et 2017 sont croisés avec la version de mai 2016 des sites Natura 2000.

Les couches utilisées sont issues du croisement des couches des sites SIC et des sites ZPS. Les géographies sont donc de trois types : les zones SIC, les zones ZPS et les zones à la fois SIC et ZPS.

Les couches Natura 2000 sont constituées de 7 colonnes. Les zones ZPS et SIC possèdent trois informations chacune : un FID, un SITECODE et un SITENAME c'est-à-dire un numéro unique, un code normalisé pour l'identification du site et un nom de site. De plus, une colonne aire en hectare est présente.

Site SIC (zone spéciale de conservation) *Site ZPS (Zone de protection spéciale)*

<i>FID_sic160</i>	<i>FID_zps160</i>
<i>SITECODE</i>	<i>SITECODE_1</i>
<i>SITENAME</i>	<i>SITENAME_1</i>

Lors de l'import des couches Natura 2000 sur PostgreSQL une erreur est apparue. Celle-ci concernait la précision de la colonne « aire_ha ». J'ai donc dû refactoriser, modifier ce champ, avec l'outil du même nom sur QGIS, afin de redéfinir sa longueur et sa précision. Le choix s'est porté sur une longueur de 18 caractères et de 12 en précision. Afin de faire ce traitement j'ai dû corriger les géométries avec l'outil « réparer les géométries ». J'ai enfin pu importer la couche créée sur PostgreSQL.

Pour un souci de performance de matériel, lié au télétravail, j'ai utilisé le RPG divisé par région et non le RPG « France entière ». Les tables ont été importées sur le logiciel avec la fonction ogr2ogr. J'ai commencé par corriger les géométries en créant une nouvelle table avec les géométries valides. (Annexe 4, script 9)

J'ai ensuite intersecté chaque région avec la couche Natura 2000. Ce sont uniquement les informations concernant les sites Natura 2000 et l'identifiant unique des parcelles (« id_parcel ») qui sont conservés. (Annexe 4, script 10)

Il existe donc quatre cas possibles :

1. La parcelle n'est ni dans une zone SIC, ni dans une zone ZPS ;
2. La parcelle est dans une zone SIC ;
3. La parcelle est dans une zone ZPS ;
4. La parcelle est à la fois dans une zone SIC et une zone ZPS.

Ainsi les parcelles qui ne sont dans aucune des deux zones, ne seront pas dans cette table ; celles qui sont sur les limites de plusieurs sites seront « découpées en plusieurs morceaux » de parcelles. Les géométries sont donc modifiées.

Il est alors nécessaire de calculer les surfaces de ces « morceaux » de parcelles et la proportion de chaque parcelle en zone Natura 2000. (Annexe 4, script 11 à 16)

<i>Etapas</i>	<i>Objectif</i>	<i>Script</i>
Ajouter une colonne « surface » et calculer la surface avec la fonction « ST_Area »	Connaître la surface de chaque parcelle en zone Natura 2000	11
Ajouter une colonne « surface_géo » et faire une jointure avec la table du RPG afin de récupérer la surface géographique de la parcelle initiale	Connaître la surface des parcelles initiales	12
Ajouter une colonne « pourc_surf » et calculer le pourcentage de surface de chaque « morceau » de parcelle par rapport à la parcelle initiale	Connaître la part de chaque parcelle en zone Natura 2000	13
Ajouter une colonne « pourc_surf_zps » et calculer le pourcentage de surface de chaque parcelle en zone ZPS	Connaître la part de chaque parcelle en zone ZPS	14
Ajouter une colonne « pourc_surf_sic » et calculer le pourcentage de surface de chaque parcelle en zone SIC	Connaître la part de chaque parcelle en zone SIC	15

Ajouter une colonne « pourc_tot » et calculer le pourcentage de surface de chaque parcelle en zone Natura 2000	Connaître la part totale de chaque parcelle en zone Natura 2000	16
--	---	----

Afin de connaître la nature des cultures présentes en zone Natura 2000, il est nécessaire de faire une jointure entre la table créée précédemment et celle du RPG. De la même manière, il est important de récupérer le code Insee de la commune sur laquelle la parcelle est située. Cela permet d'agrémenter la base de données de l'ODR qui se rapporte à l'identité « commune ».

<i>Etapes</i>	<i>Objectif</i>	<i>Script</i>
Ajouter une colonne « code Insee » et faire une jointure avec la table du RPG afin de récupérer le code Insee de la commune de la parcelle	Faire le lien entre « id_parcel » et « code_insee » ce qui permettra d'agrémenter la base de données de l'ODR	17
Ajouter une colonne « code_group » et faire une jointure avec la table du RPG afin de récupérer le code du groupe de cultures présent sur la parcelle	Connaître le groupe de la culture présente sur la parcelle	17
Ajouter une colonne « culture_d1 » et faire une jointure avec la table du RPG afin de récupérer le code du groupe de la culture dérobée 1.	Connaître le groupe de la culture dérobée 1 présente sur la parcelle	17
Ajouter une colonne « culture_d2 » et faire une jointure avec la table du RPG afin de récupérer le code du groupe de la culture dérobée 2	Connaître le groupe de la culture dérobée 2 présente sur la parcelle	17

Pour agrémenter la base de données de l'ODR il est nécessaire de concentrer toutes les tables « régions » en une unique table « France entière ». Lors de cette étape, les tables ont été « compactées » afin d'obtenir une ligne par code Insee et par groupe de culture. Les surfaces en zones « SIC » ont été sommées dans la colonne « surf_sic », de même pour les surfaces en zones « ZPS » dans la colonne « surf_zps ». Enfin, la colonne « surf_siczps » correspond à la surface totale présente en site Natura2000. Attention, celle-ci peut être différente de la somme des autres colonnes car les sites « SIC » et « ZPS » peuvent être superposés. (Annexe 4, script 18)

Comme lors du croisement RPG/communes, cette table non-géographique a été exportée au format « .csv » afin d'être traitée sur le logiciel PhpMyAdmin.

Le traitement effectué vise à transformer la table afin d'obtenir une ligne par commune. C'est alors cette table qui a été importée sur la base de données de l'ODR afin d'agrémenter la fiche du réseau « système agricole » de la plateforme concernant les proportions et les surfaces en zone Natura 2000.

3. Optimisation des scripts

Les scripts des croisements présentés précédemment ont été conservés afin d'être réutilisés pour les campagnes prochaines (2018, 2019...).

Pour faciliter l'utilisation de ces scripts par d'autres personnes, un travail d'optimisation et de description a été nécessaire.

Script de croisement du RPG avec la couche des limites communales

La première mission concerne une optimisation des étapes, notamment en diminuant le temps nécessaire pour effectuer les traitements et la taille des tables créées.

Pour diminuer les temps de traitements de nombreux index ont été ajoutés. Par exemple, l'ajout d'un index sur la colonne « surface » de la table « rpg_commune » permet de faire évoluer le temps de traitement pour le calcul de la colonne « surf_maxi » de 19 secondes à 5 secondes. Ainsi, le temps a diminué de près de 75%.

De même, l'extraction de la colonne « surf_maxi » se faisait grâce à une table temporaire. Or, après plusieurs recherches, je me suis rendu compte que ce traitement était possible sans créer cette table temporaire mais en ajoutant une condition dans le « where ». (Figure 12)

<i>Avant</i>	<i>Après</i>
<pre>alter table fanny.tb_rpg_2015_com_94_corse add column surf_maxi numeric (10,5) create temporary table tata as (select id_parcel, max (surface) as maxim from fanny.tb_rpg_2015_com_94_corse group by id_parcel); create index tataindex ON tata (id_parcel); update fanny.tb_rpg_2015_com_94_corse set surf_maxi= (select maxim from tata where tata.id_parcel = fanny.tb_rpg_2015_com_94_corse.id_parcel); Drop table tata;</pre>	<pre>alter table schema.tb_rpg_20xx_com_94_corse add column surf_maxi numeric (10,5) update schema.tb_rpg_20xx_com_94_corse set surf_maxi= (select max(surface) from schema.tb_rpg_20xx_com_94_corse as tbl where tbl.id_parcel = schema.tb_rpg_20xx_com_94_corse.id_parcel group by tbl.id_parcel);</pre>

Figure 12: Comparaison entre le premier script créé pour le calcul de "surf_maxi" et celui utilisé dans le script final

Ceci a permis un gain de temps d'environ 20% sur cette étape.

Lors du dénombrement des communes qui intersectent chaque parcelle, il était nécessaire de créer une nouvelle table. Celle-ci résultait de l'ancienne table à laquelle on ajoutait une colonne « nbr_com ». L'optimisation a permis d'ajouter cette colonne sans devoir créer une nouvelle table (Figure 13).

<i>Avant</i>	<i>Après</i>
<pre>create table fanny.tb_rpg2015_reg_76_occitanie as (WITH tb3 as (SELECT * FROM fanny.tb_rpg_2015_reg_76_occitaniev), tb4 as (select count (*) as nbr_com , tbl.id_parcel as id_parcel FROM fanny.tb_rpg_2015_reg_76_occitaniev as tbl, fanny.tb_commune_2014_metrople as tb2 where st_intersects (tbl.wkb_geometry::geometry,tb2.wkb_geometry::geometry) group by id_parcel) SELECT tb3.*, tb4.nbr_com FROM tb3 LEFT JOIN tb4 ON tb3.id_parcel=tb4.id_parcel); ALTER TABLE fanny.tb_rpg2015_reg_76_occitanie ADD CONSTRAINT pk_2015_reg_occ PRIMARY KEY (ocq_fid); ALTER TABLE fanny.tb_rpg2015_reg_76_occitanie ALTER COLUMN wkb_geometry TYPE geometry (multipolygon) USING ST_SetSRID(wkb_geometry,1184); CREATE INDEX reg_occ_2015_wkb_geometry_idx ON fanny.tb_rpg2015_reg_76_occitanie USING GIST (wkb_geometry);</pre>	<pre>alter table schema.tb_rpg_20xx_reg_11_ile_de_francev add column nbr_com integer; update schema.tb_rpg_20xx_reg_11_ile_de_francev set nbr_com =(select count (*) FROM schema.tb_rpg_20xx_reg_11_ile_de_francev as tbl, schema.tb_commune_2014_metrople where st_intersects (tbl.wkb_geometry::geometry,tb2.wkb_geometry::geometry) and tbl.id_parcel=schema.tb_rpg_20xx_reg_11_ile_de_francev.id_parcel group by id_parcel);</pre>

Figure 13: Comparaison entre le premier script créé pour le calcul de "nbr_com" et celui utilisé dans le script final

Ceci a permis un gain de temps mais surtout un gain d'espace de stockage puisque la table n'est pas doublée comme précédemment.

La deuxième mission concerne une description des étapes. En effet, le script sera réutilisé par d'autres personnes qui ne sont pas nécessairement celles qui ont suivi mon stage. Il est donc nécessaire de décrire l'ensemble des étapes et les modifications à apporter.

Les scripts utilisent le langage SQL et sont enregistrés de façon à être modifiés sur le logiciel Notepad++. Les commentaires s'effectuent avec des tirets et apparaissent de couleur verte. Il est possible de supprimer/ajouter des caractères. Afin de modifier une chaîne de caractères par une autre dans tout le script, c'est la fonction « rechercher/remplacer » qui est utilisée. Le raccourci du clavier pour l'utiliser est « Ctrl + H ». Ainsi, le nom du schéma (arborescence de la base de données sur PostgreSQL), qui est présent 615 fois dans le script, peut être remplacé en seulement quelques secondes. De même pour l'année du RPG à traiter qui est représentée par « xx » dans le script 745 fois.

La redéfinition de ces caractères est nécessaire dès le début des traitements. Ensuite, les connexions à la base de données doivent être modifiées ainsi que les chemins d'accès aux fichiers du RPG par régions.

Après ces modifications, les traitements jusqu'à l'étape de récupération des codes Insee peuvent tous être lancés via PostgreSQL en une fois. Lorsque ces traitements sont terminés il est alors possible de lancer la récupération des codes Insee région par région. S'il fonctionne alors la région suivante peut être lancée. Si le message d'erreur « plus d'une ligne renvoyée par sous-requête » est renvoyé, ça signifie qu'une parcelle a été découpée par une limite communale de manière égale. Il faut alors lancer un autre script dans lequel seuls les noms de table doivent être renseignés. Cette manipulation a concerné environ 3 tables par années traitées. Il aurait été possible de la généraliser à toutes les régions mais le temps de traitement est beaucoup plus important et une colonne est ajoutée ce qui augmente la capacité de stockage nécessaire.

Les dernières étapes du script peuvent être lancées et la table nommée « schema.tb_rpg20xx_commune_code_group » pourra être exportée au format « .csv » afin d'être traitée sur le logiciel phpMyAdmin et importée sur la base alimentant la plateforme de l'ODR.

Script du croisement du RPG avec la couche des limites des sites de *Natura 2000*

Le script de l'intersection des couches RPG et Natura 2000 a été réalisé, en partie, avant la fin du script précédent. Certaines étapes sont inutiles si le script précédent est réalisé dans un premier temps, en totalité. L'optimisation de ce script a donc reposé sur la suppression d'étapes en modifiant les tables initiales utilisées et les colonnes conservées lors des traitements.

Les premières étapes concernaient l'import des données sur le logiciel PostgreSQL et la correction des géométries invalides. Celles-ci ont été supprimées puisque déjà réalisées lors du script du croisement RPG/communes.

Lors de l'intersection des couches du RPG avec la couche des sites Natura 2000, les colonnes conservées dans la table créée sont modifiées. En effet, les colonnes « SITECODE » et « SITENAME » des sites Natura 2000 ne sont pas conservées comme cela avait été fait précédemment car elles ne sont pas utiles pour la plateforme ODR. Toutefois, si un partenaire a besoin de cette information, il sera possible de la retrouver en faisant une jointure selon le code « FID », unique à chaque site. Les colonnes concernant le code Insee, le code du groupe de culture et ceux des cultures dérobées ont été conservées ce qui évite de devoir réaliser une jointure pour les récupérer dans les étapes suivantes. Ces étapes ont donc été supprimées.

De la même manière que dans le script précédent, les étapes ont été décrites et expliquées pour l'utilisateur. Seuls le schéma et l'année sont à modifier et l'ensemble du script peut être lancé.

III. Résultats et discussion

1. Les tables de données

A l'issue de mon stage, les tables de données géographiques, résultantes des croisements effectués, pour les RPG de 2015, 2016 et 2017 seront conservées par l'ODR. Ce sont quatre lots de treize tables :

- Les tables issues du croisement géographique entre le RPG et les limites communales : « tb_rpg_20xx_com_ » + numéro et nom de la région. L'entité géographique correspond aux parcelles découpées selon les limites des communes de 2014. Elles possèdent divers attributs : deux concernant la parcelle initiale (« id_parcel » et « surf_parc »), les informations concernant la culture présente sur la parcelle (« code_cultu », « code_group », « cultude_d1 » et « cultude_d2 »), celles en lien avec la commune (« code_insee » et « nom_com ») et enfin celles avec l'entité géographique (« surface », « surf_maxi » et « id_new_parc » (issu de la contraction de l'identifiant de la parcelle et du code Insee)).

wkb_geometry geometry	id_new_parcel [PK] text	id_parcel character varying (10)	surf_parc numeric (8,2)	code_cultu character varying (3)	code_group character varying (2)	culture_d1 character varying (3)	culture_d2 character varying (3)	code_insee character varying (5)	nom_com character varying (45)	surface numeric (10,5)	surf_maxi numeric (10,5)
01060000206A0800000...	10184778_2A288	10184778	2.30	PPH	18	[null]	[null]	2A288	Sotta	2.29845	2.29845
01060000206A0800000...	10184779_2A288	10184779	0.15	BOP	17	[null]	[null]	2A288	Sotta	0.15288	0.15288
01060000206A0800000...	10184780_2A288	10184780	0.41	BOP	17	[null]	[null]	2A288	Sotta	0.41180	0.41180
01060000206A0800000...	10184781_2A288	10184781	2.87	BOP	17	[null]	[null]	2A288	Sotta	2.86618	2.86618

- Les tables du RPG initiales auxquelles plusieurs champs ont été ajoutés : « nbr_com » (nombre de commune que la parcelle intersecte), « code_insee » (code Insee de la commune où la surface de la parcelle est majoritaire) et « surface_ge » (surface de la parcelle en ha avec une précision de l'ordre du m²).

wkb_geometry geometry	ogc_fid [PK] integer	id_parcel character varying (10)	surf_parc numeric (8,2)	code_cultu character varying (3)	code_group character varying (2)	culture_d1 character varying (3)	culture_d2 character varying (3)	nbr_com bigint	code_insee character varying (6)
01060000206A0800000...	1	10184778		2.30 PPH	18	[null]	[null]	1	2A288
01060000206A0800000...	2	2951454		0.50 BOP	17	[null]	[null]	1	2A288
01060000206A0800000...	3	10184779		0.15 BOP	17	[null]	[null]	1	2A288
01060000206A0800000...	4	2951513		2.95 PPH	18	[null]	[null]	1	2A288

- Les tables du RPG avec les mêmes informations que les précédentes mais dans laquelle seules les parcelles dont le code Insee appartient à un des départements de la région étudiée sont conservées (tb_rpg_20xx_reg_dept »).
- Les tables issues du croisement géographique entre le RPG et les limites des sites Natura2000. Elles contiennent les « morceaux » de parcelles découpées selon les sites SIC et ZPS. Les parcelles qui ne sont pas sur ces sites ne sont donc pas présentes dans ces tables. Les champs sont : ceux identifiant les sites Natura 2000, ceux concernant la parcelle (identifiant, cultures...) et ceux concernant l'intersection (surface, pourcentage en site SIC et en site ZPS, pourcentage total en zone Natura 2000).

Toutes ces tables permettront de donner des informations complémentaires à un partenaire s'il en fait la demande, telles que les sources d'erreurs (parcelles hors des limites communales...) ou des précisions (site Natura 2000 précis...).

Enfin, ce sont six tables non géographiques qui ont été créées : trois correspondent à l'union, par année, des

tables présentées au troisième point ci-dessus et les trois autres à l'union de celles présentées au dernier point ci-dessus. Ces six tables viennent alimenter, après une dernière modification sur phpMyAdmin, la base de données de la plateforme ODR (figure 14).

commune	gc1	gc2	gc3	gc4	gc5	gc6	gc7	gc8	gc9	gc10	gc11	gc12	gc13	gc14
01001	201.08	279.13	67.13	0	27.93	15.47	44.46	0	0	0	18.67	0	0	0
01002	0	0	0	0	0	0	0	0	0	0	0	0	0	0
01004	18.49	8.88	43.51	0	0	6.35	0	0	0	0	8.39	0	0	0
01005	200.31	418.83	12.02	35.38	17.69	51.55	23.82	0	0	0	12.95	0	0	0
01006	0	8.72	0	1.49	0	0	0	0	0	0	0	0	0	0
01007	205.07	571.26	24.31	0	26.11	62.02	168.53	4.51	0	0	40.97	0	0	0
01008	72.62	22.11	79.11	0	35.1	1.78	0	0	0	0	4.27	0	0	0
01009	2.14	22.15	0.66	1.15	0	9.35	0	0	0	0	2.41	0	0	0

Figure 14: Vue de la table de données qui alimente la plateforme ODR (gc= Groupe de culture : gc1 = blé, gc2 = maïs, etc.)

Indicateurs de surfaces RPG (en milliers d'ha) par Départements

Note :

- Dans le tableau ci-dessous apparaissent les 28 groupes de cultures du regroupement initial ASP.
- Pour plus de détails, merci de consulter la note méthodologique de ce dossier.

code	Départements	Blé tendre	Maïs grain et ensilage	Orge	Autres céréales	Colza	Tournesol	Autres oléagineux	Proteagineux	Plantes à fibres	Semences	Gel	Gel industriel
01	Ain	34.01	65.92	8.11	5.83	6.15	3.40	5.07	0.29	0.01	0.00	3.88	0.00
02	Aisne	190.72	32.61	47.45	3.21	51.68	0.26	0.26	13.61	2.37	0.00	8.28	0.00
03	Allier	49.22	39.54	15.27	19.91	13.03	3.87	0.62	1.21	0.01	0.00	2.79	0.00
04	Alpes-de-Haute-Provence	2.76	0.96	3.20	13.20	1.80	1.89	0.56	0.78	0.03	0.00	2.25	0.00
05	Hautes-Alpes	3.00	0.69	3.19	3.32	0.11	0.17	0.01	0.04	0.00	0.00	0.05	0.00
06	Alpes-Maritimes	0.01	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
07	Ardèche	2.34	1.97	2.13	3.92	0.28	0.80	0.07	0.10	0.00	0.00	0.90	0.00
08	Ardennes	66.91	26.73	33.46	5.21	21.35	0.06	0.44	4.16	0.17	0.00	1.70	0.00
09	Ariège	8.76	10.78	2.18	5.16	1.21	5.54	0.91	0.55	0.02	0.00	1.07	0.00
10	Aube	109.10	12.23	90.30	1.44	50.29	3.84	2.95	6.57	3.80	0.00	9.16	0.00
11	Aude	6.92	2.97	5.48	35.07	2.51	21.74	0.88	2.16	0.01	0.00	10.56	0.00
12	Aveyron	14.01	15.79	24.28	27.85	1.46	0.41	0.06	0.43	0.01	0.00	0.09	0.00
13	Bouches-du-Rhône	1.72	1.18	1.38	17.00	0.68	3.33	0.67	0.21	0.01	0.00	2.95	0.00
14													
15	<												

Indicateurs de surfaces RPG en taux par Départements

Note :

- Dans le tableau ci-dessous apparaissent les 28 groupes de cultures du regroupement initial ASP.
- Les pourcentages sont calculés par rapport à la surface totale du RPG (y compris le gc28 divers).
- Pour plus de détails, merci de consulter la note méthodologique de ce dossier.

code	Départements	% Blé tendre	% Maïs grain et ensilage	% Orge	% Autres céréales	% Colza	% Tournesol	% Autres oléagineux	% Proteagineux	% Plantes à fibres	% Semences	% Gel	% Gel industriel
01	Ain	13.35	25.87	3.18	2.29	2.41	1.33	1.99	0.11	0.00	0.00	1.52	0.00
02	Aisne	38.48	6.58	9.57	0.65	10.43	0.05	0.05	2.75	0.48	0.00	1.67	0.00
03	Allier	10.15	8.15	3.15	4.11	2.69	0.80	0.13	0.25	0.00	0.00	0.58	0.00
04	Alpes-de-Haute-Provence	1.01	0.35	1.17	4.85	0.66	0.69	0.20	0.29	0.01	0.00	0.82	0.00
05	Hautes-Alpes	1.19	0.27	1.26	1.32	0.04	0.07	0.00	0.01	0.00	0.00	0.02	0.00
06	Alpes-Maritimes	0.01	0.00	0.01	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
07	Ardèche	1.71	1.44	1.56	2.87	0.20	0.58	0.05	0.08	0.00	0.00	0.66	0.00
08	Ardennes	21.64	8.65	10.82	1.68	6.91	0.02	0.14	1.35	0.05	0.00	0.55	0.00
09	Ariège	8.76	10.78	2.18	5.16	1.21	5.54	0.91	0.55	0.02	0.00	1.07	0.00
10	Aube	109.10	12.23	90.30	1.44	50.29	3.84	2.95	6.57	3.80	0.00	9.16	0.00
11	Aude	6.92	2.97	5.48	35.07	2.51	21.74	0.88	2.16	0.01	0.00	10.56	0.00
12	Aveyron	14.01	15.79	24.28	27.85	1.46	0.41	0.06	0.43	0.01	0.00	0.09	0.00
13	Bouches-du-Rhône	1.72	1.18	1.38	17.00	0.68	3.33	0.67	0.21	0.01	0.00	2.95	0.00
14													
15	<												

Figure 15: Vue d'un exemple de résultat sur la plateforme

2. Les scripts

La seconde mission de mon stage était de créer des scripts pour ajouter facilement les campagnes à venir (2018, 2019 ...) sur ODR.

Ces scripts ont été réalisés au langage SQL adapté au logiciel PostgreSQL (Postgis). Ils sont donc stockés sur un fichier compatible avec l'éditeur de script Notepad++.

Le premier script (annexe 6) propose le croisement géographique du RPG avec les limites communales. Les étapes recouvrent tous les traitements : de l'import des fichiers nécessaires sur la base de données à l'export de ceux créés. Des commentaires décrivent tout le long les étapes et les manipulations à effectuer par l'utilisateur.

Il en est de même pour le second script (annexe 7) qui correspond au croisement géographique du RPG avec les sites Natura 2000.

Les deux scripts sont liés. En effet, il est nécessaire de les effectuer dans le sens ci-dessus car le second utilise les tables créées pendant les traitements du premier script.

3. La note méthodologique

J'ai aussi effectué une note méthodologique pour mon maître de stage qui décrit, de manière synthétique, l'ensemble des données utilisées dans le cadre de la préparation des tables de 2015 à 2017, les traitements effectués et les cas particuliers relevés. (Annexe 8)

Concrètement, pour un partenaire de l'ODR, il est aujourd'hui possible de consulter sur la plateforme web les tableaux d'indicateurs créés (Figure 15), de les exporter en fonction de l'échelle souhaitée afin de réaliser des statistiques et donc d'évaluer les politiques agricoles mises en place (RPG, Natura 2000, indicateurs de diversités, prairies...).

IV. Conclusion

1. Bilan pour l'entreprise

J'ai pu répondre à l'attente principale de mon entreprise d'accueil qui était de proposer des agrégations communales du registre parcellaire graphique afin d'alimenter la plateforme web de l'unité. Cette dernière disposait jusqu'à là des données antérieures à 2015. J'ai donc préparé les agrégations du RPG pour les années 2015 à 2017, données que l'ODR possédait. J'ai ensuite préparé un script au langage SQL qui permettra d'effectuer les mêmes traitements pour les années à venir.

Ma deuxième mission était d'agrémenter les données du RPG avec des données d'intérêts écologiques. Il avait été question d'étudier deux types de données : celles des sites Natura 2000 et celles de la couche Végétation de la Bd-Topo de l'IGN. A cause du télétravail, lié à l'épidémie de Covid-19, et des moyens informatiques qu'il engendre, il n'a pas été possible de traiter la couche de Végétation de la BD-Topo. Je me suis donc concentrée davantage sur les sites Natura 2000 en effectuant les croisements géographiques du RPG avec ces sites. De même que précédemment, j'ai créé un script SQL qui permettra d'effectuer ces traitements pour les années suivantes.

Les tables créées ont déjà permis d'agrémenter plusieurs fiches de la plateforme web de l'ODR.

2. Bilan personnel

Je tire un bilan positif de ce stage puisqu'il m'a été bénéfique pour plusieurs raisons.

Pour commencer, j'ai pu mettre en pratique de nombreuses connaissances apprises dans le cadre de ma licence. En effet, j'ai beaucoup utilisé les logiciels de traitement de données géographiques : PostgreSQL et QGIS.

Ce stage m'a permis d'acquérir de nouvelles connaissances et compétences. J'ai approfondi un sujet qui m'intéressait beaucoup car, à travers le RPG, il établissait un lien entre les pratiques agricoles et les politiques mises en œuvre pour les encadrer. J'ai aussi développé progressivement et tout au long de mon stage mes compétences dans la gestion de base de données et dans le langage SQL. Le contexte particulier de la Covid-19 m'a de fait placé en situation d'autonomie totale m'obligeant à organiser seule mon travail et à prendre les initiatives nécessaires pour résoudre les problèmes et difficultés auxquels je me trouvais confrontée.

J'ai par exemple pris conscience de la nécessité d'une optimisation des scripts en apprenant, au fil du stage, l'existence et les possibilités que pouvaient donner de nouvelles fonctions.

Bibliographie

- Livret présentation INRA « Le centre de recherche de Toulouse Midi-Pyrénées » : D. Charamel, S. Chaveroux, J. Juille, Septembre 2014.
- « Livret d'accueil US-ODR INRAE » : Janvier 2020
- « Dossier d'évaluation US-ODR : présentation de la plateforme ODR »
- Le cahier des techniques de l'INRA : GPS & SIG Pour la conduite de dispositifs expérimentaux Vers l'émergence de la communauté métier « GéoExpé » : INRA, 2014
- Document : Réseau Système Agricole, V2.3 P. Cantelaube
- « Gestion et disponibilité des données du RPG sur la plateforme logicielle ODR » : P. Cantelaube INRA / ODR, Janvier 2012
- « Note méthodologique : Dossier thématique 'Indicateurs (exploitations et cultures) RPG version de base' » : P. Cantelaube INRA / ODR, nov. 2013
- « Projet 'Systèmes de cultures et systèmes d'exploitations' : Note Méthodologique » : P. Cantelaube, nov. 2015
- <https://www.ecologique-solidaire.gouv.fr/reseau-europeen-natura-2000-1>
- https://geoservices.ign.fr/ressources_documentaires/Espace_documentaire/BASES_VECTORIELLES/RPG/DC_DL_RPG_2-0.pdf

Table des figures

Figure 1: Extrait du Registre Parcellaire Graphique de 2015	1
Figure 2: Organigramme de l'US-ODR Source: ODR.....	6
Figure 3: Schéma représentant le cheminement des données du Réseau Système Agricole. Source : ODR : Réseau Système Agricole, V2.3 P. Cantelaube).....	7
Figure 4: Tableau des différents niveaux d'anonymisation du RPG. Source: ODR.....	8
Figure 5: Tableau des projections du RPG. Source: IGN.....	8
Figure 6: Données attributaires du RPG version 2.0	8
Figure 7: Vue de la plateforme web ODR.....	9
Figure 8: Carte de comparaison de la résolution des deux sources de limites communales, ainsi que des limites des parcelles RPG.....	11
Figure 9: Carte de comparaison des limites communales sur le RPG de 2015.	12
Figure 10: Carte de comparaison des parcelles du RPG 2017 (à gauche : Occitanie au 1 ^{er} plan et Nouvelle-Aquitaine au 2 nd plan ; à droite l'inverse).....	17
Figure 11: Vue d'une partie de la dernière table fournie.....	17
Figure 12: Comparaison entre le premier script créé pour le calcul de "surf_maxi" et celui utilisé dans le script final	21
Figure 13: Comparaison entre le premier script créé pour le calcul de "nbr_com" et celui utilisé dans le script final	21
Figure 14: Vue de la table de données qui alimente la plateforme ODR	23
Figure 15: Vue d'un exemple de résultat sur la plateforme.....	24

Table des sigles

ASP : Agence de Services et Paiements

CMS : Content Management System

DOM-TOM : Départements et Territoires d'Outre-Mer

IGN : Institut National de l'Information Géographique et forestière

INRA : Institut National de la Recherche Agronomique

INRAE : Institut National de la Recherche pour l'Agriculture, l'alimentation et l'environnement

INSEE : Institut national de la statistique et des études économiques

IRIS : Ilots Regroupés pour l'Information Statistique

IRSTEA : Institut National de Recherche en Science et Technologies pour l'Environnement et l'Agriculture

MAE : Mesures Agro-Environnementales

ODR : Observatoire du Développement Rural

PAC : Politique Agricole Commune

RPG : Registre Parcellaire Graphique

RSA : Réseau Systèmes Agricoles

SAU : Surface Agricole Utile

SGBD : Système de Gestion de Base de Données

SIC : zones spéciales de conservation dans le cadre du programme Natura2000

SIG : Système d'Informations Géographiques

SQL : Structured Query Language

SRID : Spatial Reference Identifier

ZPS : Zones de Protection Spéciales pour les oiseaux dans le cadre du programme Natura2000

Annexes

1. Tableau de correspondance des groupes de cultures

<i>Groupes de culture</i>	
Code groupe	Libellé
1	Blé tendre
2	Maïs grain et ensilage
3	Orge
4	Autres céréales (avoine, blé dur...)
5	Colza
6	Tournesol
7	Autres oléagineux
8	Protéagineux
9	Plantes à fibres
11	Gel (surfaces gelées, sans production)
14	Riz
15	Légumineuses à grains
16	Fourrage
17	Estives et landes
18	Prairies permanentes
19	Prairies temporaires
20	Vergers
21	Vignes
22	Fruits à coque
23	Oliviers
24	Autres cultures industrielles
25	Légumes ou fleurs
26	Canne à sucre
28	Divers

2. Carte des régions Françaises et codes associés



Codes des régions

Code	Nom de la région
11	Ile-de-France
24	Centre – Val de Loire
27	Bourgogne-Franche-Comté
28	Normandie
32	Hauts-de-France
44	Grand Est
52	Pays de la Loire
53	Bretagne
75	Nouvelle-Aquitaine
76	Occitanie
84	Auvergne-Rhône-Alpes
93	Provence-Alpes-Côte d'Azur
94	Corse

3. Tableau des erreurs de géométries constatées sur le RPG

<i>Régions</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
<i>Ile de France</i>	0	19	0
<i>Centre val de Loire</i>	0	70	0
<i>Bourgogne Franche Comté</i>	3	>70	0
<i>Normandie</i>	5	>70	0
<i>Hauts de France</i>	6	>70	0
<i>Grand-Est</i>	8	>70	0
<i>Pays de la Loire</i>	13	>70	0
<i>Bretagne</i>	1	>70	0
<i>Nouvelle aquitaine</i>	3	>70	0
<i>Occitanie</i>	5	>70	2
<i>Auvergne Rhône Alpes</i>	8	>70	0
<i>Pacca</i>	0	>70	0
<i>Corse</i>	1	36	1

4. Scripts SQL

Script 1 :

```
create table fanny.tb_rpg_2015_reg_27_bourgogne_franche_comtev as (  
  select ST_MakeValid (wkb_geometry) as wkb_geometry,  
  ogc_fid, id_parcel, surf_parce, code_cultu, code_group, culture_d1, culture_d2  
  from fanny.tb_rpg_2015_reg_27_bourgogne_franche_comte) ;  
  
ALTER TABLE fanny.tb_rpg_2015_reg_27_bourgogne_franche_comtev ADD CONSTRAINT pk_bfcv PRIMARY KEY (ogc_fid);  
ALTER TABLE fanny.tb_rpg_2015_reg_27_bourgogne_franche_comtev  
  ALTER COLUMN wkb_geometry TYPE geometry (multipolygon, 2154) USING ST_SetSRID(wkb_geometry,2154);  
create index bfc_wkb_geometry_idx ON fanny.tb_rpg_2015_reg_27_bourgogne_franche_comtev USING GIST (wkb_geometry);
```

Script 2 :

```
select st_intersection (tbl1.wkb_geometry::geometry,tb2.wkb_geometry::geometry) as geom,  
  tbl1.id_parcell as id_parcell,tbl1.surf_parce as surf_parce, tb2.depcom as depcom  
  
FROM fanny.tb_rpg_2015_reg_bourgogne_franche_comte as tbl1, fanny.tb_commune_2014_metropole as tb2  
  
where st_intersects(tbl1.wkb_geometry::geometry,tb2.wkb_geometry::geometry))
```

Script 3 :

```
create table fanny.tb_rpg_2015_com_ll_ile_de_france as (  
  select st_multi(st_collectionExtract(st_intersection (tbl1.wkb_geometry::geometry,tb2.wkb_geometry::geometry),3)) as wkb_geometry,  
  concat(id_parcel,'_',depcom) as id_new_parcel, id_parcel, surf_parce, code_cultu, code_group, culture_d1, culture_d2, depcom as code_insee, nom_com  
  FROM fanny.tb_rpg_2015_reg_ll_ile_de_france as tbl1, fanny.tb_commune_2014_metropole as tb2  
  where st_intersects(tbl1.wkb_geometry::geometry,tb2.wkb_geometry::geometry)  
  );  
ALTER TABLE fanny.tb_rpg_2015_com_ll_ile_de_france ADD CONSTRAINT pk_2015_com_idf PRIMARY KEY (id_new_parcel);  
ALTER TABLE fanny.tb_rpg_2015_com_ll_ile_de_france ALTER COLUMN wkb_geometry TYPE geometry (multipolygon, 2154) USING ST_SetSRID(wkb_geometry,2154);  
create index com_idf_2015_wkb_geometry_idx ON fanny.tb_rpg_2015_com_ll_ile_de_france USING GIST (wkb_geometry);
```

Script 4 :

```
# Ajouter une colonne avec calcul d aire:  
  
#créer la colonne  
  
ALTER TABLE fanny.tb_rpg_2015_com_bourgogne_franche_comte ADD COLUMN  
surface NUMERIC (10, 5)  
  
# mettre à jour la colonne avec la valeur d aire en ha  
  
UPDATE fanny.tb_rpg_2015_com_bourgogne_franche_comte  
set surface = (ST_Area(geom)/10000)
```

Script 5 :

```
alter table rpg2016.tb_rpg_2016_com_ll_ile_de_france add column surf_maxi NUMERIC (10, 5);  
create temporary table tatall as (  
  select id_parcel, max (surface) as maxim  
  from rpg2016.tb_rpg_2016_com_ll_ile_de_france  
  group by id_parcel  
  );  
create index tatallindex ON tatall (id_parcel);  
update rpg2016.tb_rpg_2016_com_ll_ile_de_france  
  set surf_maxi= (  
    select maxim  
    from tatall  
    where tatall.id_parcel = rpg2016.tb_rpg_2016_com_ll_ile_de_france.id_parcel  
  );  
Drop table tatall;
```


Script 10 :

```
create table natura2000.tb_rpg_2016_nat_ll_ile_de_france as (  
SELECT rpg.id_parcel, nat.fid_sic160, nat.sitecode, nat.sitename, nat.fid_zps160, nat.sitecode_l, nat.sitename_l,  
st_multi(st_collectionExtract(st_intersection (rpg.wkb_geometry::geometry,nat.wkb_geometry::geometry),3)) as wkb_geometry  
FROM natura2000.tb_natura2000_2016 as nat, natura2000.tb_rpg_2016_reg_ll_ile_de_francev as rpg  
WHERE st_intersects(rpg.wkb_geometry::geometry,nat.wkb_geometry::geometry)  
);  
Alter table natura2000.tb_rpg_2016_nat_ll_ile_de_france add column id SERIAL PRIMARY KEY;  
ALTER TABLE natura2000.tb_rpg_2016_nat_ll_ile_de_france  
ALTER COLUMN wkb_geometry TYPE geometry (multipolygon, 2154) USING ST_SetSRID(wkb_geometry,2154);  
create index nat_idf_2016_wkb_geometry_idx  
ON natura2000.tb_rpg_2016_nat_ll_ile_de_france USING GIST (wkb_geometry);
```

Script 11 :

```
ALTER TABLE fanny.tb_rpg_2015_nat_94_corse ADD COLUMN surface NUMERIC (10, 5);  
UPDATE fanny.tb_rpg_2015_nat_94_corse set surface = (ST_Area(wkb_geometry)/10000);
```

Script 12 :

```
ALTER TABLE fanny.tb_rpg_2015_nat_ll_ile_de_france add COLUMN surface_geo numeric (10,5) ;  
UPDATE fanny.tb_rpg_2015_nat_ll_ile_de_france  
SET surface_geo =  
    (select a.surface_geo  
    from rpg2015.tb_rpg2015_reg_ll_ile_de_france as a  
    where a.id_parcel = fanny.tb_rpg_2015_nat_ll_ile_de_france.id_parcel)  
;
```

Script 13 :

```
ALTER TABLE fanny.tb_rpg_2015_nat_94_corse ADD COLUMN pourc_surf NUMERIC (10, 2);  
UPDATE fanny.tb_rpg_2015_nat_94_corse  
SET pourc_surf =  
    ((surface/surface_geo)*100)  
;
```

Script 14 :

```
create index nat_idf_2015_id_parcel_idx ON fanny.tb_rpg_2015_nat_ll_ile_de_france (id_parcel);  
ALTER TABLE fanny.tb_rpg_2015_nat_ll_ile_de_france add COLUMN pourc_surf_zps numeric (10,2) ;  
UPDATE fanny.tb_rpg_2015_nat_ll_ile_de_france  
SET pourc_surf_zps = (  
    select sum (a.pourc_surf)  
    from fanny.tb_rpg_2015_nat_ll_ile_de_france as a  
    where a.fid_zps160 != '-1' and a.id_parcel= fanny.tb_rpg_2015_nat_ll_ile_de_france.id_parcel  
    group by a.id_parcel  
);
```

Script 15 :

```
ALTER TABLE fanny.tb_rpg_2015_nat_94_corse add COLUMN pourc_surf_sic numeric (10,2) ;  
UPDATE fanny.tb_rpg_2015_nat_94_corse  
SET pourc_surf_sic = (  
    select sum (a.pourc_surf)  
    from fanny.tb_rpg_2015_nat_94_corse as a  
    where a.fid_sic160 != '-1' and a.id_parcel= fanny.tb_rpg_2015_nat_94_corse.id_parcel  
    group by a.id_parcel  
);
```

Script 16 :

```
ALTER TABLE fanny.tb_rpg_2015_nat_ll_ile_de_france add COLUMN pourc_tot numeric (10,2) ;
UPDATE fanny.tb_rpg_2015_nat_ll_ile_de_france
SET pourc_tot = (
    select sum (a.pourc_surf)
    from fanny.tb_rpg_2015_nat_ll_ile_de_france as a
    where a.id_parcel= fanny.tb_rpg_2015_nat_ll_ile_de_france.id_parcel
    group by a.id_parcel
);
```

Script 17 :

```
ALTER TABLE fanny.tb_rpg_2015_nat_ll_ile_de_france add COLUMN code_insee character varying (6);
ALTER TABLE fanny.tb_rpg_2015_nat_ll_ile_de_france add COLUMN code_group character varying (2);
ALTER TABLE fanny.tb_rpg_2015_nat_ll_ile_de_france add COLUMN culture_d1 character varying (3);
ALTER TABLE fanny.tb_rpg_2015_nat_ll_ile_de_france add COLUMN culture_d2 character varying (3);
UPDATE fanny.tb_rpg_2015_nat_ll_ile_de_france
SET code_insee = (
    select code_insee
    from fanny.tb_rpg2015_reg_dept_ll_ile_de_france as a
    where a.id_parcel= fanny.tb_rpg_2015_nat_ll_ile_de_france.id_parcel
);
UPDATE fanny.tb_rpg_2015_nat_ll_ile_de_france
SET code_group = (
    select code_group
    from fanny.tb_rpg2015_reg_dept_ll_ile_de_france as a
    where a.id_parcel= fanny.tb_rpg_2015_nat_ll_ile_de_france.id_parcel
);
UPDATE fanny.tb_rpg_2015_nat_ll_ile_de_france
SET culture_d1 = (
    select culture_d1
    from fanny.tb_rpg2015_reg_dept_ll_ile_de_france as a
    where a.id_parcel= fanny.tb_rpg_2015_nat_ll_ile_de_france.id_parcel
);
UPDATE fanny.tb_rpg_2015_nat_ll_ile_de_france
SET culture_d2 = (
    select culture_d2
    from fanny.tb_rpg2015_reg_dept_ll_ile_de_france as a
    where a.id_parcel= fanny.tb_rpg_2015_nat_ll_ile_de_france.id_parcel
);
create index nat_idf_2015_code_insee_idx ON fanny.tb_rpg_2015_nat_ll_ile_de_france (code_insee);
create index nat_idf_2015_code_group_idx ON fanny.tb_rpg_2015_nat_ll_ile_de_france (code_group);
```

Script 18 :

```
create table fanny.tb_rpg_2015_nat_commune_code_group as (
(select code_insee, code_group, culture_d1, culture_d2,
    sum(a.surface * (case when a.fid_sic140 != '-1' then 1
    else 0 end) ) as surf_sic,
    sum(a.surface * (case when a.fid_zps140 != '-1' then 1
    else 0 end) ) as surf_zps,
    sum(a.surface) as surf_siczps
from fanny.tb_rpg_2015_nat_ll_ile_de_france as a
group by code_insee, code_group, culture_d1, culture_d2
order by code_insee)
union all
(select code_insee, code_group, culture_d1, culture_d2,
    sum(a.surface * (case when a.fid_sic140 != '-1' then 1
    else 0 end) ) as surf_sic,
    sum(a.surface * (case when a.fid_zps140 != '-1' then 1
    else 0 end) ) as surf_zps,
    sum(a.surface) as surf_siczps
from fanny.tb_rpg_2015_nat_24_centre_val_de_loire as a
group by code_insee, code_group, culture_d1, culture_d2
order by code_insee)
union all
```

5. Résultats du croisement géographique de RPG avec les limites communales

<i>Régions</i>	<i>2015</i>		<i>2016</i>		<i>2017</i>	
	Parcelles non-intersectées	Parcelles divisées en deux surfaces égales	Parcelles non-intersectées	Parcelles divisées en deux surfaces égales	Parcelles non-intersectées	Parcelles divisées en deux surfaces égales
Ile de France	0	0	0	0	0	0
Centre val de Loire	0	0	0	0	0	0
Bourgogne Franche Comté	0	0	0	0	0	0
Normandie	47	0	46	0	46	0
Hauts de France	72	1	70	2	64	1
Grand-Est	46	0	42	0	43	1
Pays de la Loire	20	0	21	0	21	0
Bretagne	287	0	275	0	298	0
Nouvelle aquitaine	63	2	70	0	74	0
Occitanie	8	3	8	2	8	2
Auvergne Rhône Alpes	17	0	16	0	16	0
Pacca	0	0	21	1	20	1
Corse	0	0	0	0	0	0
Total	560	6	569	5	590	5

6. Script SQL du croisement géographique RPG/ Communes (extrait)

```
---SCRIPT POUR LE CROISEMENT DES TABLES DU RPG AVEC CELLE DES COMMUNES APIN DE RECUPERER LE CODE INSEE POUR CHAQUE PARCELLE---

--#Enregistrer sous le fichier afin de garder la version originale
--#modifier le nom du schéma (arborescence de la base de données) utilisé pour tout le script: 'Ctrl'+ 'H' : recherche: 'schema.' ; remplacer par le nom du schema '+'
--#pareil pour l année à traiter: recherche: 'xx' ; remplacer par l année

-----
---importer les fichiers shapex de la table 'commune' et des tables du RPG par région (France entière trop lourd comme traitement) sur postgis avec l invite de commande
-----

---ouvrir PgAdmin et QGIS---

---connecter la base de données PostgreSQL à QGIS: sur QGIS: base de données: gestionnaire de base de données: Postgis---

---ouvrir l invite de commande de QGIS OSGeo 4W Shell---

    ---importer la table 'commune'---
        --#modifier les connexions à la base de données ('Ctrl'+ 'H')et le chemin d accès à la table. Pour le nom de la table le modifier avec 'Ctrl'+ 'H' car réutilisé
ogr2ogr -f "PostgreSQL" PG:"host=localhost user=postgres dbname=rpg password=ggat" -nlt MULTIPOLYGON -a_srs EPSG:2154 -s_srs EPSG:2154
E:/communes2014/tb_commune_2014_metropole.shp -nln schema.tb_commune_2014_metropole

    ---importer les tables du RPG par régions---
        --#modifier les connexions à la base de données ('Ctrl'+ 'H') et les chemins d accès aux tables
ogr2ogr -f "PostgreSQL" PG:"host=localhost user=postgres dbname=rpg password=ggat" -nlt MULTIPOLYGON -a_srs EPSG:2154 -s_srs EPSG:2154
E:/rpg20xx/RPG_2-0_SHP_LAMB94_R11-20xx/PARCELLES_GRAPHIQUES.shp -nln schema.tb_rpg_20xx_reg_11_ile_de_france

-----
---Corriger les géométries invalides, définir la clé primaire et le SRID et ajouter des indexes---
-----

create index com_depcom_idx ON schema.tb_commune_2014_metropole (depcom);

create table schema.tb_rpg_20xx_reg_11_ile_de_france as (
    select ST_MakeValid (wkb_geometry) as wkb_geometry,
           ogr_fid, id_parcel, surf_parc, code_cultu, code_group, culture_d1, culture_d2
    from schema.tb_rpg_20xx_reg_11_ile_de_france );
ALTER TABLE schema.tb_rpg_20xx_reg_11_ile_de_france ADD CONSTRAINT pk_idfux PRIMARY KEY (ogr_fid);
ALTER TABLE schema.tb_rpg_20xx_reg_11_ile_de_france ALTER COLUMN wkb_geometry TYPE geometry (multipolygon, 2154) USING ST_SetSRID(wkb_geometry,2154);
create index idfux_wkb_geometry_idx ON schema.tb_rpg_20xx_reg_11_ile_de_france USING GIST (wkb_geometry);
create index idfux_id_parcel_idx ON schema.tb_rpg_20xx_reg_11_ile_de_france (id_parcel);

-----
---Supprimer les tables exportées au début---
-----

drop table schema.tb_rpg_20xx_reg_11_ile_de_france;

-----
---Intersecter les régions avec la tables des communes: créer une table, définir PK et indexes---
-----

create table schema.tb_rpg_20xx_com_11_ile_de_france as (
select st_multi(st_collectionExtract(st_intersection (tbl1.wkb_geometry::geometry,tb2.wkb_geometry::geometry),3)) as wkb_geometry,
        concat(id_parcel,'_',depcom) as id_new_parcel, id_parcel, surf_parc, code_cultu, code_group, culture_d1, culture_d2, depcom as code_insee, nom_com
FROM schema.tb_rpg_20xx_reg_11_ile_de_france as tbl1, schema.tb_commune_2014_metropole as tb2
where st_intersects(tbl1.wkb_geometry::geometry,tb2.wkb_geometry::geometry)
);
ALTER TABLE schema.tb_rpg_20xx_com_11_ile_de_france ADD CONSTRAINT pk_20xx_com_idf PRIMARY KEY (id_new_parcel);
ALTER TABLE schema.tb_rpg_20xx_com_11_ile_de_france ALTER COLUMN wkb_geometry TYPE geometry (multipolygon, 2154) USING ST_SetSRID(wkb_geometry,2154);
create index com_idf_20xx_wkb_geometry_idx ON schema.tb_rpg_20xx_com_11_ile_de_france USING GIST (wkb_geometry);
CREATE INDE/ com_idf_20xx_parcel_idx ON schema.tb_rpg_20xx_com_11_ile_de_france (id_parcel);

-----
---Ajouter une colonne surface---
-----

ALTER TABLE schema.tb_rpg_20xx_com_11_ile_de_france
    ADD COLUMN surface NUMERIC (10,5);
UPDATE schema.tb_rpg_20xx_com_11_ile_de_france
    set surface = (ST_Area(wkb_geometry)/10000);
CREATE INDE/ com_idf_20xx_surface_idx ON schema.tb_rpg_20xx_com_11_ile_de_france (surface);
```

```

-----Ajouter une colonne surf_maxi qui contient la surface de la plus grande partie de la parcelle après découpage selon les communes-----
alter table schema.tb_rpg_20xx_com_ll_ile_de_france add column surf_maxi numeric (10,5);
update schema.tb_rpg_20xx_com_ll_ile_de_france
set surf_maxi= (
select max (surface)
from schema.tb_rpg_20xx_com_ll_ile_de_france as tbl
where tbl.id_parcel = schema.tb_rpg_20xx_com_ll_ile_de_france.id_parcel
group by tbl.id_parcel
);

-----Ajouter une colonne qui compte le nombre de communes que la parcelle intersecte-----
alter table schema.tb_rpg_20xx_reg_ll_ile_de_france add column nbr_com integer;
update schema.tb_rpg_20xx_reg_ll_ile_de_france set nbr_com =(
select count (*)
FROM schema.tb_rpg_20xx_reg_ll_ile_de_france as tbl, schema.tb_commune_2014_metropole as tb2
where st_intersects (tbl.wkb_geometry::geometry,tb2.wkb_geometry::geometry) and tbl.id_parcel=schema.tb_rpg_20xx_reg_ll_ile_de_france.id_parcel
group by id_parcel
);

-----Ajouter une colonne avec le code_insee de la commune sur laquelle la surface de la parcelle est majoritaire-----

--Lancer les traitements un par un: si erreur "plus d une ligne renvoyée par sous requete" (cela correspond aux parcelles découpées de manière égale) .
ALTER TABLE schema.tb_rpg_20xx_reg_ll_ile_de_france ADD COLUMN code_insee character varying (6);
update schema.tb_rpg_20xx_reg_ll_ile_de_france set code_insee =
(select a.code_insee
from schema.tb_rpg_20xx_com_ll_ile_de_france as a
where a.id_parcel = schema.tb_rpg_20xx_reg_ll_ile_de_france.id_parcel and a.surface= a.surf_maxi );

---Cas particuliers: pour les erreurs: "plus d une ligne renvoyée par sous requete" ---
#remplacer le nom du schema,le nom de la table du RPG commune : 'tablecom' (Ctrl + H) et le nom de la table du RPG region : 'tablereg' (Ctrl + H )

--ajout d une colonne nb_surf_mm qui compte le nombre de morceau de parcelle qui ont la meme surface
Alter table schema.tablecom add column nb_surf_mm numeric (5,1);
update schema.tablecom set nb_surf_mm =
(select count (distinct a.code_insee)
from schema.tablecom as a
where a.surface= a.surf_maxi and a.id_parcel=schema.tablecom.id_parcel);

---completer la colonne code_insee pour les cas normaux: nb_surf_mm = 1 ---
ALTER TABLE schema.tablereg ADD COLUMN code_insee character varying (6);
update schema.tablereg set code_insee =
(select a.code_insee
from schema.tablecom as a
where a.nb_surf_mm=1 and a.id_parcel = schema.tablereg.id_parcel and a.surface= a.surf_maxi );

---completer la colonne code_insee pour les cas où nb_surf_mm > 1 ---
update schema.tablereg set code_insee = (
WITH tb3 as (
SELECT tbl.id_parcel
FROM schema.tablereg as tbl),
tb4 as (
select min(code_insee) as codinsee, a.id_parcel as parcel, a.nb_surf_mm
from schema.tablecom as a
where nb_surf_mm>1
group by a.id_parcel, a.nb_surf_mm
)
SELECT tb4.codinsee
FROM tb3
LEFT JOIN tb4 ON tb3.id_parcel=tb4.parcel
where tb4.nb_surf_mm>1 and tb3.id_parcel=schema.tablereg.id_parcel
)where code_insee is null;

```

```

-----Ajouter une colonne surface avec la superficie en ha des parcelles de base pour comparer avec surf_parc-----
-----

ALTER TABLE schema.tb_rpg_20m_reg_ll_ile_de_francev
  ADD COLUMN surface_ge NUMERIC (10, 5);
UPDATE schema.tb_rpg_20m_reg_ll_ile_de_francev
  set surface_ge = (ST_Area(wkb_geometry)/10000);

-----
---Ne garder par table que les parcelles qui sont sur un des département de la région: supprimer celles qui sont sur le département voisin---
-----
---bien conserver ces tables: elles seront utilisées pour le croisement RPG/Natura2000---

create table schema.tb_rpg20m_reg_dept_ll_ile_de_france as (
select *
from schema.tb_rpg_20m_reg_ll_ile_de_francev
where code_insee like '75%' or code_insee like '77%' or code_insee like '78%' or code_insee like '91%' or code_insee like '92%' or code_insee like '93%' or code_i
);
CREATE INDEX reg_dept_idf_20m_parcel_idx ON schema.tb_rpg20m_reg_dept_ll_ile_de_france (id_parcel);

-----
---créer une table, France entiere, avec la surface en ha par code culture et par commune---
-----

create table schema.tb_rpg20m_commune_code_group as (

(select code_insee, code_group, sum(surf_parc) as surf_ha
from schema.tb_rpg20m_reg_dept_ll_ile_de_france
group by code_insee, code_group
order by code_insee, code_group)

union all
...

--#exporter cette table au format .csv depuis pgadmin (choisir le chemin et le nom de la table, codage = UTF8 et cocher header) puis la transferer sur phpmyadmin

```

7. Script SQL du croisement géographique RPG/ Natura2000 (extrait)

```
---SCRIPT POUR LE CROISEMENT DES TABLES DU RPG AVEC NATURA 2000 AFIN DE CONNAITRE LES SUPERFICIES DE PARCELLES EN ZONE SIC ET EN ZONE EPS---

---À modifier dans l ordre!!!---

--#modifier le nom du schéma (arborescence de la base de données) utilisé pour tout le script: 'Ctrl'+ 'H' : recherche: 'schema.' ; remplacer par le nom du schema '+'
--#pareil pour l année de la couche natura2000 à traiter: recherche: 'natura2000_20xx' ; remplacer par 'natura2000_20' + l année de la couche natura2000
--#pareil pour l année du RPG à traiter: recherche: '20xx' ; remplacer par l année du RPG
--#pareil pour les noms des colonnes 'fid' de la table natura2000: recherche: 'fid_sic140'; remplacer par le nom de la colonne et pareil pour 'fid_sps140'

-----
---importer les fichiers shapex de la table 'natura2000' et des tables du RPG par région (France entière trop lourd comme traitement) sur postgis
-----

---ouvrir PgAdmin et QGIS---

---connecter la base de données PostgreSQL à QGIS: sur QGIS: base de données: gestionnaire de base de données: Postgis---

---ouvrir l invite de commande de QGIS OSGeo 4W Shell---

---importer la table 'natura2000'---

--#modifier les connexions à la base de données ('Ctrl'+ 'H') et le chemin d accès à la table
ogr2ogr -f "PostgreSQL" PG:"host=localhost user=postgres dbname=rpg password=ggat" -nlt MULTIPOLYGON -a_srs EPSG:2154 -s_srs EPSG:2154
E:/Natura2000/sic_U_sps_2009.shp -nln schema.tb_natura2000_20xx

---tables rpg régions à réutiliser: tb_rpg_20xx_reg_dept

-----
---Intersecter les couches rpg/natura2000---
-----

create table schema.tb_rpg_20xx_nat_l1_ile_de_france as (
  SELECT rpg.id_parcel, rpg.code_insee, rpg.code_group, rpg.culture_d1, rpg.culture_d2, rpg.surface_ge, nat.fid_sic140, nat.fid_sps140,
  st_multi(st_collectionExtract(st_intersection (rpg.wkb_geometry::geometry,nat.wkb_geometry::geometry),3)) as wkb_geometry
  FROM schema.tb_natura2000_20xx as nat, schema.tb_rpg20xx_reg_dept_l1_ile_de_france as rpg
  WHERE st_intersects(rpg.wkb_geometry::geometry,nat.wkb_geometry::geometry)
);
ALTER table schema.tb_rpg_20xx_nat_l1_ile_de_france add column id SERIAL PRIMARY KEY;
ALTER TABLE schema.tb_rpg_20xx_nat_l1_ile_de_france ALTER COLUMN wkb_geometry TYPE geometry (multipolygon, 2154) USING ST_SetSRID(wkb_geometry,2154);
create index nat_idf_20xx_wkb_geometry_idx ON schema.tb_rpg_20xx_nat_l1_ile_de_france USING GIST (wkb_geometry);

-----
---Calculer l aire de l intersection---
-----

ALTER TABLE schema.tb_rpg_20xx_nat_l1_ile_de_france ADD COLUMN surface NUMERIC (10, 5);
UPDATE schema.tb_rpg_20xx_nat_l1_ile_de_france set surface = (ST_Area(wkb_geometry)/10000);

-----
---Calculer la proportion de la surface natura 2000 par raport à la surface totale---
-----

ALTER TABLE schema.tb_rpg_20xx_nat_l1_ile_de_france ADD COLUMN pourc_surf NUMERIC (10, 2);
UPDATE schema.tb_rpg_20xx_nat_l1_ile_de_france
SET pourc_surf =
  ((surface/surface_ge)*100)
;

-----
---Calculer la proportion de surface par parcelle en sone ZPS---
-----

create index nat_idf_20xx_id_parcel_idx ON schema.tb_rpg_20xx_nat_l1_ile_de_france (id_parcel);
ALTER TABLE schema.tb_rpg_20xx_nat_l1_ile_de_france add COLUMN pourc_surf_sps numeric (10,2) ;
UPDATE schema.tb_rpg_20xx_nat_l1_ile_de_france
SET pourc_surf_sps = (
  select sum (a.pourc_surf)
  from schema.tb_rpg_20xx_nat_l1_ile_de_france as a
  where a.fid_sps140 != '-1' and a.id_parcel= schema.tb_rpg_20xx_nat_l1_ile_de_france.id_parcel
  group by a.id_parcel
);
```

---Calculer la proportion de surface par parcelle en sone SIC---

```
ALTER TABLE schema.tb_rpg_20xx_nat_ll_ile_de_france add COLUMN pourc_surf_sic numeric (10,2) ;
UPDATE schema.tb_rpg_20xx_nat_ll_ile_de_france
SET pourc_surf_sic = (
select sum (a.pourc_surf)
from schema.tb_rpg_20xx_nat_ll_ile_de_france as a
where a.fid_sic140 != '-1' and a.id_parcel= schema.tb_rpg_20xx_nat_ll_ile_de_france.id_parcel
group by a.id_parcel
);
```

---Calculer la proportion de la parcelle en sone natura 2000---

```
ALTER TABLE schema.tb_rpg_20xx_nat_ll_ile_de_france add COLUMN pourc_tot numeric (10,2) ;
UPDATE schema.tb_rpg_20xx_nat_ll_ile_de_france
SET pourc_tot = (
select sum (a.pourc_surf)
from schema.tb_rpg_20xx_nat_ll_ile_de_france as a
where a.id_parcel= schema.tb_rpg_20xx_nat_ll_ile_de_france.id_parcel
group by a.id_parcel
);
```

---creer une table avec une ligne par code group et par commune---

```
create table fanny.tb_rpg_2015_nat_commune_code_group as (
(select code_insee, code_group, culture_d1, culture_d2,
sum(a.surface * (case when a.fid_sic140 != '-1' then 1
else 0 end) ) as surf_sic,
sum(a.surface * (case when a.fid_sps140 != '-1' then 1
else 0 end) ) as surf_sps,
sum(a.surface) as surf_sicps
from fanny.tb_rpg_2015_nat_ll_ile_de_france as a
group by code_insee, code_group, culture_d1, culture_d2
order by code_insee)
union all
```

Croisement du RPG avec les couches communes et Natura 2000

Introduction

Dans le cadre de mon stage, qui a pour but d'agrémenter la base de données ODR, j'ai croisé les couches du RPG avec celles des communes afin d'obtenir le code Insee de la commune où elle est située. Ensuite les RPG ont été croisés avec la couche des sites Natura 2000 afin d'enrichir les données et de permettre d'améliorer les travaux d'évaluation et de modélisation du système agricole.

Présentation des données

Le Registre Parcellaire Graphique (RPG)

Le RPG est un système d'informations géographiques permettant l'identification des parcelles agricoles sur le Territoire français. Il repose sur un système de déclarations graphiques effectuées par les agriculteurs dans le cadre de la politique agricole commune (PAC). La version 2.0 du RPG est élaborée par l'IGN (Institut National de l'Information Géographique et forestière). L'information disponible concerne donc les parcelles des exploitations faisant l'objet d'une aide de la PAC. Les données sont fournies tous les ans, sous formes géographiques vectorielles (format shapefile ESRI), avec des données attributaires associées.

Attributs	Définition	Caractéristiques
Id_parcel	Identifiant unique de la parcelle	Caractères (max 10)
Code_cultu	Code culture principale	Caractères (3), présent en annexe
Code_group	Code groupe de la culture principale	Caractères (2), présent en annexe
Culture_d1	Code culture dérobée 1	Caractères (3), vide si nul
Culture_d2	Code culture dérobée 2	Caractères (3), vide si nul
Surf_parc	Surface en hectare de la parcelle	Décimal (8,2)

Donnée attributaire du RPG version 2.0 (à partir de 2015)

	id_parcel	code_cultu	code_group	culture_d1	culture_d2
1	0106300204030000		17		

L'« id_parcel » est conservé d'une année sur l'autre pour les années postérieures à 2015, lorsque la géométrie de la parcelle n'est pas modifiée.

Les données sont proposées dans les systèmes de référence suivants :

Zone	Système géodésique	Ellipsoïde associé	Projection	Unité	Résolution
France métropolitaine	RGF93	IAG GRS 1980	Lambert 93	m	cm
Guadeloupe, Saint-Barthélemy, Saint-Martin, Martinique	WGS84	IAG GRS 1980	UTM Nord fuseau 20	m	cm
Guyane	RGF95	IAG GRS 1980	UTM Nord fuseau 22	m	cm
La Réunion	RGR92	IAG GRS 1980	UTM Sud fuseau 40	m	cm
Mayotte	RGM04	IAG GRS 1980	UTM Sud fuseau 38	m	cm

Source :

https://geoservices.ign.fr/ressources_documentaires/Espace_documentaire/BASES_VECTORIELLES/RPG/DC_DL_RPG_2-0.pdf

Dans le cadre du travail réalisé, ce sont les RPG de 2015, 2016 et 2017 qui sont utilisés.

Les limites communales

Après comparaison entre plusieurs sources de données, c'est la couche Iris de l'INSEE (Institut National de la Statistique et des Etudes Economique) qui a été choisie. Cette couche est un découpage du territoire en mailles de taille homogène. Les mailles sont regroupées par commune et sont définies par le code Insee et le nom de la commune. L'échelle est de 1 : 1 000 000, c'est-à-dire de 10 km pour un cm. C'est la couche de 2014 qui a été utilisée. Elle a été réalisée avec les résultats du recensement de la population 2011. (Source INSEE)

	ogc_fid	depcom	nom_com	wkb_geometry
	[PK] integer	character varying (3)	character varying (45)	geometry
1	1	01002	L'Abergement-de-Varey	01060000206A0800000...

Natura 2000

Le réseau européen Natura 2000 regroupe les outils de la politique européenne de préservation de la biodiversité. Il est constitué d'un ensemble de sites naturels, terrestres et marins qui ont pour objectif de préserver la diversité biologique et le patrimoine naturel.

En France, 12,9 % de la surface terrestre métropolitaine, soit 7 millions d'hectares, est concernée par ces sites. Ils sont divisés en deux principales catégories : les zones de protection spéciales pour les oiseaux (ZPS) et les zones spéciales de conservation (ZSC ou SIC).

La couche du RPG 2015 est croisée avec la version de septembre 2014 des sites Natura 2000. Alors que les RPG 2016 et 2017 sont croisés avec la version de mai 2016 des sites Natura 2000.

Chaque couche Natura 2000 possède à la fois les données des zones SIC et celles des zones ZPS : elle a été obtenue lors du croisement des couches de chacune de ces deux zones. Ces zones sont définies par les champs suivants :

	Site SIC (zone spéciale de conservation)	Site ZPS (Zone de protection spéciale)
Identifiant unique	FID_sic160	FID_zps160
Code identifiant du site	SITECODE	SITECODE_1
Nom du site	SITENAME	SITENAME_1

La surface de chaque zone est présente dans la colonne « aire_ha ».

ogc_fid	fid_sic160	sitecode	sitename	fid_zps160	sitecode_1	sitename_1	aire_ha	wkb_geometry
[PK] integer	numero (16)	character varying (3)	character varying (190)	numero (16)	character varying (3)	character varying (190)	numero (16,12)	geometry
1	0	FR1100799	Massif de Fontainebleau	-1	FR11		4.992981474439	01060000206A0800000...

Croisement RPG & Natura 2000

Les tables du RPG résultant du croisement RPG/Communes ont été croisées avec la couche Natura 2000 afin de connaître la proportion de parcelles dans des zones de protection. Pour cela, c'est la fonction SQL « ST_intersection » qui a été utilisée. Le résultat de cette intersection est une table qui contient toutes les « parties » de parcelle qui sont sur au moins un site Natura 2000. Les parcelles qui ne sont pas sur un site ne sont pas conservées et celles qui sont sur une limite de sites sont modifiées : c'est uniquement la géométrie de la partie sur le site qui est conservée. Un identifiant unique a donc été ajouté. Il se nomme « id ».

La surface de ces parties a été calculée avec une précision de 0.1m², elle est présente dans la colonne « surface » et est en hectare.

La surface géographique (« surface_ge ») avait été calculée dans le croisement précédent. Elle correspond à la surface de l'identité « parcelle » du RPG. Elle a permis de calculer le pourcentage de la surface de la parcelle présente sur un site Natura 2000.

Pour cela, la première étape est de calculer le pourcentage de surface de chaque partie découpée par rapport à la surface géographique de la parcelle initiale. Cette information est unique pour chaque partie découpée et est présente dans la colonne « pourc_surf ».

Ensuite, pour chaque parcelle du RPG, les pourcentages de surface de chaque partie de parcelle ont été ajoutés. Le pourcentage de la surface initiale de la parcelle qui est sur une zone SIC est dans la colonne « pourc_surf_sic ». De même pour les zones ZPS, cette information est présente dans la colonne « pourc_surf_zps ». Ces informations sont donc uniques pour chaque « id_parcel ». Ainsi elles sont présentes autant de fois dans la table que la parcelle initiale a été découpée lors de l'intersection avec la couche Natura 2000. Enfin, « pourc_tot » correspond au pourcentage de surface de chaque parcelle en zone Natura 2000.

id_parcel	surface	surface_ge	pourc_surf	pourc_surf_sic	pourc_surf_zps	pourc_tot
varchar(10)	float(10,2)	float(10,2)	float(10,2)	float(10,2)	float(10,2)	float(10,2)

Pour agréementer la base de données de l'ODR il est nécessaire de concentrer toutes les tables « régions » en une unique table « France entière ». Lors de cette étape, les tables ont été « compactées » afin d'obtenir une ligne par code Insee et par groupe de culture. Les surfaces en zones « SIC » ont été sommées dans la colonne « surf_sic », de même pour les surfaces en zones « ZPS » dans la colonne « surf_zps ». Enfin, la colonne « surf_siczps » correspond à la surface totale présente en site Natura2000. Attention, celle-ci peut être différente de la somme des autres colonnes car les sites « SIC » et « ZPS » peuvent être superposés.

Cas particuliers

- La parcelle n°10185637 est présente dans la couche du RPG métropole mais pas dans couche du RPG de la région Corse (où elle se situe).
- Tableau des erreurs de géométries constatées sur le RPG :

Régions	2015	2016	2017
Ile de France	0	19	0
Centre val de Loire	0	70	0
Bourgogne Franche Comté	3	>70	0
Normandie	5	>70	0
Hauts de France	6	>70	0
Grand-Est	8	>70	0
Pays de la Loire	13	>70	0
Bretagne	1	>70	0
Nouvelle aquitaine	3	>70	0
Occitanie	5	>70	2
Auvergne Rhône Alpes	8	>70	0
Pacca	0	>70	0
Corse	1	36	1

- Résultat du croisement géographique de RPG avec les limites communales :

Régions	2015		2016		2017	
	Parcelles non-intersectées	Parcelles divisées en deux surfaces égales	Parcelles non-intersectées	Parcelles divisées en deux surfaces égales	Parcelles non-intersectées	Parcelles divisées en deux surfaces égales
Ile de France	0	0	0	0	0	0
Centre val de Loire	0	0	0	0	0	0
Bourgogne Franche Comté	0	0	0	0	0	0
Normandie	47	0	46	0	46	0
Hauts de France	72	1	70	2	64	1
Grand-Est	46	0	42	0	43	1
Pays de la Loire	20	0	21	0	21	0
Bretagne	287	0	275	0	298	0
Nouvelle aquitaine	63	2	70	0	74	0
Occitanie	8	3	8	2	8	2
Auvergne Rhône Alpes	17	0	16	0	16	0
Pacca	0	0	21	1	20	1
Corse	0	0	0	0	0	0
Total	560	6	569	5	590	5