



HAL
open science

Vers une meilleure diffusion des séquences de cultures obtenues avec le RPG : synergies entre RPG Explorer et les travaux menés à l'INRA ODR

Philippe Martin, Thomas Poméon

► To cite this version:

Philippe Martin, Thomas Poméon. Vers une meilleure diffusion des séquences de cultures obtenues avec le RPG : synergies entre RPG Explorer et les travaux menés à l'INRA ODR. [Rapport Technique] Office National de l'Eau et des Milieux Aquatiques. 2018, 15 p. hal-02437583

HAL Id: hal-02437583

<https://hal.science/hal-02437583v1>

Submitted on 13 Jan 2020

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

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



Vers une meilleure diffusion des séquences de cultures obtenues avec le RPG : synergie entre RPG Explorer et les travaux menés à l'INRA ODR

Rapport final

**Philippe MARTIN (AgroParisTech UMR SADAPT),
Thomas POMEON (INRA US ODR),**

Décembre 2018]

- **AUTEURS**

Philippe MARTIN, Professeur (AgroParisTech), philippe.martin@agroparistech.fr

Thomas POMEON, Ingénieur de Recherches (INRA), thomas.pomeon@inra.fr

- **CONTRIBUTEURS**

Pierre CANTELAUBE, Ingénieur d'Etudes (INRA), Pierre.Cantelaube@inra.fr

Nicolas PISKIEWICZ, Développeur (entreprise Piskiewicz), npiskiewicz@gmail.com

- **CORRESPONDANTS**

Partenaire : Philippe MARTIN, Professeur (AgroParisTech), philippe.martin@agroparistech.fr

Agence française pour la biodiversité : Claire BILLY, chargée de mission « pression/impact pollutions diffuses agricoles » (AFB), claire.billy@afbiodiversite.fr

Droits d'usage : accès libre

Niveau géographique : national

Couverture géographique : France

Niveau de lecture : professionnel et experts

[Le **résumé**, en une page max., est intégré dans la notice qui référence le document dans les bases documentaires qui sont mises à disposition des lecteurs]

- **TITRE, AUTEURS**

- **RESUME**

Les données RPG (Registre Parcellaire Graphique) permettent de produire des cartographies de séquences de cultures utiles pour diagnostiquer les pressions sur la ressource en eau. L'INRA ODR en lien avec l'UMR AGIR d'une part, l'UMR SADAPT d'autre part, ont mis en place des systèmes différents pour l'établissement et la mise à disposition des séquences de cultures pour les gestionnaires de ressources en eau. L'UMR AGIR et l'ODR ont établi les séquences sur l'ensemble du territoire national de 2006 à 2014 alors que l'UMR SADAPT a développé un outil (RPG Explorer) qui permet d'établir ces séquences pour tout utilisateur disposant des données RPG. Le changement de nature des données livrées à compter de 2015 (parcelle et non plus îlot) ne permet plus à l'ODR d'effectuer le calcul des séquences pour les années postérieures à 2014. De son côté RPG Explorer est un outil puissant mais tous les gestionnaires n'ont pas accès aux données RPG permettant de l'utiliser.

Un travail antérieur a permis de rapprocher les algorithmes développés par chacune des équipes pour l'établissement des séquences. Ce travail a permis d'améliorer le fonctionnement de RPG Explorer. Une intégration plus forte est toutefois envisagée dans ce rapport. Il s'agirait d'établir les séquences avec RPG Explorer, de les stocker sur un serveur à l'ODR pour permettre à des utilisateurs distants de les télécharger pour pouvoir effectuer des traitements complémentaires dans RPG Explorer, sans que l'utilisateur n'ait besoin de récupérer, en amont, les données RPG. Cette approche ne permettrait pas de mobiliser tous les modules de RPG Explorer car certains d'entre eux nécessitent les données brutes du RPG. Elle permettrait toutefois le calcul d'assolements de rotations ou l'analyse des séquences (rotation cadre).

Les utilisateurs de RPG Explorer ont un accès libre et gratuit aux données du RPG pour les années 2013 et suivantes. Notons que les données pour 2013-2014 sont tronquées (on ne donne que l'occupation majoritaire de l'îlot) mais les découpages en îlot est le bon. Les données postérieures à 2014 sont non tronquées (on reste sur l'occupation majoritaire mais il n'y a qu'une occupation du sol par parcelle). Il n'est donc pas possible d'établir les séquences avec les données tronquées de 2013-2014, ni pour les données antérieures qui ne sont pas directement disponibles. L'ODR pourra ainsi transmettre des séquences associées au numéro d'îlots de l'année de référence choisie par l'utilisateur (la plus récente de la séquence) sans autre information spatiale (simple fichier csv). L'utilisateur devra ensuite intégrer les données issues du RPG IGN dans RPG Explorer pour permettre le lien entre les séquences téléchargées depuis le serveur de l'ODR et la localisation des îlots. Les numéros d'îlots dans les données distribuées par l'IGN sont différents des numéros d'îlots utilisés avec les données RPG de référence utilisées par l'INRA. On propose de se caler sur les numéros d'îlots des données RPG mise en ligne par l'IGN pour les années 2013 et suivante.

La formulation des demandes de données se fera à partir d'un module à développer dans RPG Explorer. Les demandes peuvent se faire sur une zone spécifique ou sur l'ensemble d'un département. Les périodes pour lesquelles on pourra établir des séquences seront modulables en proposant toute période démarrant en 2006 et ayant pour dernière année depuis 2007 jusqu'à l'année la plus récente des données. Les données de séquences transmises par un lien de téléchargement pourront être intégrées dans RPG Explorer et permettre les traitements complémentaires (assolement de rotation et établissement de rotation cadre).

Du côté de l'ODR, l'idée serait d'implanter RPG Explorer sur une machine windows et que cette instance de RPG Explorer puisse se connecter à la base postgresql d'un serveur où seraient disponibles les données RPG pour le traitement des séquences et les tables de séquences résultant de ce traitement. On devrait aussi gérer la réception automatique des demandes de données ainsi que le stockage du résultat dans un espace dédié pour le téléchargement. Pour éviter une congestion des services de l'ODR mais aussi pour utiliser la même version de postgresql que RPG explorer (qui n'utilise pas nécessairement la même que les serveurs de l'ODR) il serait souhaitable d'avoir un serveur dédié à ce service. Ce service ne pourra toutefois pas se mettre en place sans moyens financiers ni main d'œuvre dédiés.

- **MOTS CLES (THEMATIQUE ET GEOGRAPHIQUE) : RPG, SEQUENCES DE CULTURES, BASE DE DONNEES, POSTGRESQL, SERVEUR**

SOMMAIRE

1.1. Rappel des protocoles propre à chaque approche	5
1.1.1. RPG Explorer.....	5
1.1.2. ODR-AGIR.....	5
1.2. Des défis majeurs à relever pour les deux approches	5
1.3. Un premier rapprochement sur les algorithmes... ..	6
1.3.1. Dans un contexte de croissance des besoins en données par les utilisateurs finaux...6	
1.3.2. Qui nécessite une intégration plus poussée et un programme à moyen terme.....	6
2. Rappels sur le fonctionnement de RPG Explorer	7
2.1. Retour sur les séquences de cultures et leur valorisation dans RPG Explorer	7
2.2. Les modules de RPG Explorer non utilisables sans autres éléments que la table de séquences.....	8
3. Conséquences et contraintes au niveau de l'ODR.....	9
3.1. Les données minimales à avoir pour utiliser les séquences en provenance de l'ODR	9
3.2. Choix de la période sur laquelle sont établies les séquences de groupes cultures	9
3.3. Entre traitement anticipé ou traitement au fil de l'eau à l'ODR... ..	10
3.4. Entre envoi en direct et envoi différé des données de séquence.....	10
4. Création d'un module spécifique dans RPG Explorer	10
4.1. Demande standardisée.....	11
4.2. Intégration dans la base locale des données issues de l'ODR.....	11
5. Modification de la base de l'ODR pour une ouverture vers RPG Explorer ..	11
6. Glossaire.....	14
7. Sigles & Abréviations	15

Introduction

Ce rapport initie le développement d'une plus grande interopérabilité informatique entre les outils et données des équipes AGIR-ODR d'un côté et SADAPT de l'autre. A ce stade, l'idée est de poser les bases d'un système qui puisse permettre aux utilisateurs de RPG Explorer d'interroger la base de l'ODR pour télécharger (gratuitement) les données RPG prétraitées (séquences de groupes culture) et les utiliser dans des protocoles de traitement complémentaires (passage des numéros de groupes cultures au nom des cultures, établissement de rotations cadre, voire d'assolements de rotation).

1.1. Rappel des protocoles propres à chaque approche

1.1.1. RPG Explorer

Actuellement les utilisateurs de RPG Explorer intègrent les données RPG dans une base de données locale générée sur leur ordinateur lors de l'installation de l'outil. Les données RPG utilisées sont celles fournies par l'ASP (niveau 4) de 2006 à 2014. Les données délivrées à la parcelle depuis 2015 sont aussi compatibles avec RPG Explorer¹. L'utilisateur doit donc au préalable avoir acquis les données (payantes pour 2006-2014, gratuites pour les années suivantes). Elles doivent être sur un dossier local (répertoire du disque dur, clé USB) pour pouvoir les intégrer dans la base de données de RPG Explorer. Les séquences sont établies pour les îlots d'intersection des couches RPG de chacune des années considérées. Ces îlots d'intersection ont pour identifiant unique la concaténation des numéros d'îlots pris par la même entité physique sur la période considérée dans les données RPG

1.1.2. ODR-AGIR

Du côté de l'ODR, les séquences ont été établies sur la période 2006-2014 pour la France entière. Les données de séquences de groupes cultures sont établies pour la couche d'îlot RPG la plus récente (2014). Les données sont stockées dans une base localisée à Toulouse. Lorsque l'ODR reçoit une demande de données sur les séquences, l'extraction se fait à Toulouse par le personnel de l'ODR. Elles sont ensuite transmises par lien de téléchargement aux structures qui en ont fait la demande contre une participation aux frais de gestion de la base si le demandeur n'est pas un laboratoire INRA ou affilié. L'extraction est un fichier csv qui associe à chaque identifiant d'îlot 2014 une ou plusieurs séquence de groupes cultures.

1.2. Des défis majeurs à relever pour les deux approches

Les développeurs des deux approches sont confrontés à des défis qui les conduisent à envisager de fortes convergences.

RPG Explorer est un outil totalement paramétrable pour lequel on développe de nouveaux modules chaque année. Toutefois l'accumulation de nombreux modules rend l'interface complexe pour des usagers ne l'utilisant pas fréquemment; cela malgré une documentation de l'outil importante sous forme écrite ou vidéo. Il en résulte que l'outil malgré ses améliorations continues reste actuellement sous utilisé alors que les utilisateurs potentiels saisissent de mieux en mieux l'enjeu des données qu'il peut fournir (voir rapport sur le bilan sur l'utilisation de RPG Explorer). Par ailleurs le financement du développement de RPG Explorer n'est actuellement assuré que jusqu'à 2019 alors que le champ des développements possibles est loin d'être totalement couvert.

La méthode ODR-AGIR a permis de produire les données de séquences de cultures pour la France entière et pour la période 2006-2014. Ce travail important a nécessité la collaboration essentielle de l'UMR AGIR et de l'ODR (P. Casel, O. Théron, avec la participation de V. Fuzeau et R. Chartier). Or, suite au départ du chercheur à l'origine des développements sur la RPG (O. Théron) l'UMR AGIR ne souhaite plus travailler sur l'établissement des séquences. De son côté l'ODR n'a pas non plus les moyens en personnel pour reprendre la méthode développée par l'UMR AGIR et l'appliquer aux données RPG récentes (2015 et suivantes), qui de surcroît sont dans un format différent (information à la parcelle et pas seulement à l'îlot, passage de 28 à 309 classes de groupes de cultures). Il sera donc très difficile voire impossible à l'ODR d'établir les

¹ A la date où nous rédigeons ce document les données disponibles n'intègrent pas les informations (identifiant d'exploitation) nécessaires pour reconstituer les territoires d'exploitation pour les années postérieures à 2014. Les données d'occupation du sol sont quant à elles en libre accès avec une meilleure résolution que les données précédentes ASP niveau 4.

séquences de groupes cultures pour les années 2015 et suivantes et donc de les mettre à disposition des utilisateurs finaux. Par ailleurs, actuellement, l'extraction des données de séquences à la demande se fait contre rétribution car cela demande du temps et l'ODR n'a pas de financements spécifiques pour ce type de travail, ce qui freine aussi la diffusion des résultats.

Cette situation était en partie connue lors du dépôt du projet à l'origine de ce rapport mais elle s'est renforcée au fil du temps. L'établissement des séquences de culture constitue une ressource transversale stratégique pour la recherche et l'expertise, mais il est difficile de financer de manière pérenne et stable ces opérations (financement anticipé et récurrent d'un outil générique vs. financement de projet de recherche précis pour une période donnée, etc.).

1.3. Un premier rapprochement sur les algorithmes...

Les premiers échanges entre les deux équipes de développement ont montré l'intérêt d'une convergence. Cette convergence a été initiée par l'analyse des résultats générés par leurs algorithmes respectifs. Cette analyse s'est accompagnée d'une évolution des algorithmes de RPG Explorer pour tenir compte des apports spécifiques de la méthode ODR-AGIR. Les évolutions ont aussi eu pour origine la correction de bugs communs aux deux approches ou encore la mise en œuvre d'idées nouvelles issues de la réflexion commune. A noter que si l'algorithme de RPG Explorer a pu évoluer à l'issue de ce premier travail, cela n'a pas été le cas pour l'algorithme AGIR pour lequel, comme indiqué précédemment, on ne dispose plus de compétences permettant ni de le modifier ni de le mettre en œuvre en l'état

1.3.1. Dans un contexte de croissance des besoins en données par les utilisateurs finaux...

Dans le même temps l'utilisation des données de type « séquences de cultures » s'est développée en France sous l'impulsion de différents programmes de recherche. Le programme ABC'Terre financé par l'ADEME ambitionne la mise en place d'une chaîne de traitement de données permettant de mieux caractériser les émissions de gaz à effet de serre d'origine agricole. Cela passe par une meilleure intégration de la répartition spatiale des systèmes de culture qui s'appuie sur la connaissance des séquences de culture telle établie à partir du RPG. Le programme en est à sa deuxième phase ABC'Terre 2A (2016-2019) qui vise à transférer la méthodologie à des animateurs locaux. On s'appuie ici sur l'utilisation de RPG Explorer car l'objectif est d'aller jusqu'aux assolements de rotation caractérisant un territoire. Les formations à RPG Explorer que nous avons organisées pour ces animateurs de territoires pilotes ont montré la difficulté que pouvaient avoir ce type d'acteurs de formation agronomique à entrer dans la logique de traitement de données spatialisées portée par RPG Explorer quand ils n'ont aucune base en système d'information géographique. Même si RPG Explorer ne nécessite que peu de connaissances en ce domaine, l'absence totale de connaissance sur le sujet constitue un plafond de verre qui gêne fortement la compréhension des étapes de traitement ainsi que l'interprétation des sorties.

Le programme MAELIA développé par O. Théron maintenant à Colmar vise aussi une utilisation des séquences de groupes de cultures pour modéliser l'usage d'un territoire agricole. Ces séquences sont supposées connues et disponibles sur l'étendue nationale. L'idée sous-jacente étant de s'appuyer sur l'établissement des séquences telles qu'établies par le passé par AGIR. Nous avons vu précédemment que cette disponibilité à l'échelon national était fortement remise en question en l'état actuel des choses.

1.3.2. Qui nécessite une intégration plus poussée et un programme à moyen terme.....

Les éléments qui précèdent conduisent à proposer un mode de fonctionnement beaucoup plus intégré entre les deux approches. Il est ainsi envisagé qu'à terme une version adaptée de RPG Explorer soit implantée à l'ODR pour permettre la reconstitution des séquences de cultures mais aussi potentiellement d'autres sorties proposées par l'outil. Ces sorties seraient stockées à l'ODR et seraient accessibles à distance par des utilisateurs de RPG Explorer qui utiliseraient ce logiciel pour interroger les bases de données de l'ODR, rendant ainsi inutile le traitement en local pour ce type de données (qui implique le téléchargement des couches RPG et la manipulation de RPG-Explorer pour réaliser les étapes successives nécessaires).

Il s'agit là d'un projet à moyen terme qui implique des développements non négligeables. En effet le développement de RPG Explorer a jusqu'alors été orienté vers des utilisateurs ayant peu de bases en SIG et bases de données ce qui n'est pas le cas des ingénieurs de l'ODR. Il a aussi été optimisé pour des traitements sur quelques départements en même temps. La version de RPG Explorer à mettre en place à l'ODR, bien que s'appuyant sur le même moteur informatique, nécessiterait d'être optimisé pour permettre un traitement séquentiel de l'ensemble des départements français. Les bases de données nécessaires (couches géographiques et attribut des îlots RPG) seraient hébergées sur l'ODR, comme c'est déjà le cas, et un module RPG-Explorer pourrait s'y connecter directement (au lieu de le faire en local). L'ODR utilise déjà le système de gestion de base de données relationnelles et objet PostGreSQL, également employé par RPG-Explorer. De plus, l'ODR dispose d'un accès privilégié à des outils permettant de traiter et gérer des gros volumes de données (serveur, cœur de calcul, etc.). Le principal frein consiste à traduire en code exécutable (dans un ou plusieurs langages à préciser) les opérations faites avec RPG-Explorer, alors que ce dernier se base pour l'instant sur des objets spécifiques et des commandes faites par l'utilisateur sur l'interface RPG-Explorer. Ce développement, nécessaire si on souhaite bénéficier à nouveau d'une couverture nationale des séquences de culture telle qu'elle était précédemment mise en place et pérenniser la maintenance et le développement de RPG-Explorer au sein d'un organisme public (le travail de développement est actuellement sous-traité à un prestataire), nécessiterait des crédits non disponibles actuellement.

Dans le cadre de ce premier projet relatif au rapprochement des méthodes et des résultats nous ne traiterons que d'un volet plus réduit consistant à réfléchir aux grandes lignes d'un cahier des charges de développements informatiques permettant à RPG Explorer d'interroger la base de l'ODR pour récupérer les séquences de groupes cultures actuellement ou prochainement disponibles afin d'intégrer ces données dans les chaînes de traitement proposées par le logiciel.

2. Rappels sur le fonctionnement de RPG Explorer

Dans la mesure où la finalité est d'intégrer les données ODR dans RPG Explorer, il nous paraît important de revenir un peu plus dans le détail sur la manière dont ces données de séquences sont actuellement produites dans RPG Explorer et les défis techniques que cela pourrait poser si on s'appuyait uniquement sur l'accès aux données ODR.

2.1. Retour sur les séquences de cultures et leur valorisation dans RPG Explorer

Dans RPG Explorer, comme évoqué au début de ce rapport, on doit d'abord intégrer les données RPG, puis faire l'intersection entre les couches d'îlots des différentes années prises en compte. C'est ensuite à cette couche d'intersection préalablement nettoyée du « bruit »² que l'on va affecter les séquences de groupes cultures. La détermination des séquences n'est qu'une étape dans le protocole de traitement des données offert par RPG Explorer. Ces données de séquences pourront ensuite être utilisées seules ou en association avec d'autres issues elles aussi de traitement des données RPG pour produire de nouvelles connaissances. Nous détaillons ci-dessous les différents traitements possibles.

1. Exportation simple des données de séquences pour traitement externe à RPG Explorer : RPG Explorer permet d'exporter les données de séquences sous deux formats, un format *shape* ou un format de données tabulaires (*csv*). Le *shape* exporté se base sur la couche d'intersection des îlots. On a une ligne par îlot avec, pour donnée attributaire, une chaîne de texte décrivant la ou les séquences présentes dans l'îlot. On peut aussi se focaliser sur certains secteurs particuliers d'un département ou d'un assemblage de départements³ pour l'extraction des

² Ce bruit est constitué par des îlots d'intersection qui résultent d'une imprécision géographique dans le tracé des îlots. Ces îlots n'ont pas d'existence réelle, on les supprime pour qu'ils ne perturbent pas la détermination des séquences de groupes cultures dans les « vrais » îlots d'intersection

³ Rappelons que RPG Explorer exploite les bases RPG fournies par l'ASP et que ces bases étaient par département jusqu'en 2014. Les données disponibles à partir de 2015 sont maintenant par région, toutefois l'intégration de ces données dans RPG Explorer est faite en ventilant les données régionales en autant de bases

- données (e.g. une aire d'alimentation de captage). Dans le fichier tabulaire exporté on a une ligne par séquence donc potentiellement plusieurs lignes pour un même îlot d'intersection.
2. Conversion des séquences de groupes cultures en séquences de cultures : ce module de RPG Explorer permet à partir d'une table de séquences de groupes cultures (de 1 à 28) de générer une table de séquences de cultures plus détaillées. La conversion se fait soit sur la base d'un lien unique entre un numéro de groupe culture et un nom de culture (par exemple le groupe de cultures 6 correspond uniquement au tournesol) soit en fréquence. Ainsi en lieu et place du « groupe 2 » on peut inscrire « Orge » dans tous les cas ou soit « Orge de Printemps » ou « Orge d'hiver » en fonction de fréquences utilisées en termes de probabilité d'apparition dans les îlots (par exemple 20% d'orge de printemps et 80% d'orge d'hiver). Cette probabilité d'apparition peut venir d'une expertise locale ou de statistiques agricoles à différents niveaux ('le plus souvent départementales).
 3. Analyse des séquences et rotations cadres : partant d'une table de séquences, RPG Explorer en fait une première analyse globale (surface et fréquences des séquences identifiées) puis y repère des régularités pour ramener chaque séquence à une possible rotation. Ce module agrège aussi l'information pour rassembler dans une même catégorie différentes cultures afin de faciliter la lecture des rotations cadres (ex : blé et orge sont associés dans une même catégorie « céréales »). Dans cette approche, chaque îlot conserve une information spécifique propre en termes de rotation. C'est ce type d'approche qui est utilisé également dans la plateforme MAELIA développée par O. Théron et dans une série de données produites à partir des séquences reconstituée par AGIR.
 4. Détermination d'assolements de rotations : partant d'une table de séquences, RPG Explorer détermine un assolement de rotations constitué d'un nombre limité de rotations. Dans ce module les rotations sont exprimées en cultures et pas en format englobant de type « céréales » comme pour les rotations cadre. Dans cette approche tous les îlots d'une même zone sur laquelle on a déterminé l'assolement de rotation portent en sortie la même information, à savoir un assolement de rotation complet. C'est ce type d'approche qui est utilisé dans la méthode ABC'Terre. Les assolements de rotation sont établis sur des zones pouvant être des entités de gestion écologique (Aire d'Alimentation de Captage), des types de sol ou des territoires d'exploitation de même type.

2.2. Les modules de RPG Explorer non utilisables sans autres éléments que la table de séquences

A côté des modules que nous avons rappelés ci-dessus, il existe un certain nombre d'autres fonctionnalités qui ne fonctionnent pas à partir d'une table de séquences mais à partir des données de base des RPG annuels ou même à partir de la table d'intersection des îlots réalisée en amont de l'établissement des séquences. Dans l'hypothèse où l'utilisateur n'aurait pas les données RPG de base mais uniquement la table de séquences en provenance de l'ODR, ces modules ne pourraient pas être activés. Nous les listons ci-dessous pour rappel :

- **Module indicateur sur données ASP** : ce module permet de calculer toute une batterie d'indicateurs à partir des données annuelles du RPG sans intersection des années entre elles. On va ainsi pouvoir calculer des assolements annuels sur des espaces écologiques ou sur des territoires d'exploitation ou le croisement des deux. On peut aussi calculer des indicateurs de structure de parcellaire (dispersion des îlots dans les exploitations).
- **Module exploitation** : ce module s'appuie à nouveau sur les données RPG de base. Il en fait une intersection et intègre à la table résultante l'information sur les numéros anonyme d'exploitation associés à chaque îlot d'intersection pour chacune des années prises en compte. Cette procédure permet ensuite de déterminer des dynamique de parcellaire d'exploitation (agrandissement réduction) tant au niveau collectif (dynamique globale sur un territoire) qu'individuel (fiche d'évolution du parcellaire d'une exploitation donnée).

A noter qu'en plus des modules non activables, la seule mise à disposition d'une table de séquences sans autres informations limitera certaines options des modules qui s'appuie sur ce type de tables. C'est ainsi qu'il ne sera pas possible d'établir des assolements de rotation pour tout découpage spatial impliquant les types d'exploitation vu que l'information sur les exploitations ne sera pas présente dans les tables de séquences de groupes cultures. Par ailleurs L'ODR produit et diffuse une série d'indicateurs construits à partir du RPG à différentes échelles

(administratives en particuliers), dont certains rapprochent de ceux proposés par RPG-Explorer : indicateurs de diversité des assolements, etc.

3. Conséquences et contraintes au niveau de l'ODR

Partant de la connaissance du fonctionnement de RPG Explorer nous allons maintenant analyser les implications sur les données à fournir par l'ODR.

3.1. Les données minimales à avoir pour utiliser les séquences en provenance de l'ODR

Actuellement l'ODR transmet aux utilisateurs finaux les données de séquences sous forme d'un fichier tabulaire comparable à la sortie Excel permise par RPG Explorer (cf. point 1 sur la valorisation des séquences). Aucune information spatiale n'est contenue dans ce fichier autre que l'identifiant des îlots de l'année la plus récente auxquels sont associés les séquences. Or l'utilisation de ces données dans RPG Explorer peut nécessiter une information spatiale liée aux îlots pour les traitements suivants (assolements de rotation sur une zone particulière, cf. point 4 sur la valorisation des séquences).

La mise en ligne sur le site de l'IGN des données RPG niveau 2⁴, utilisables librement par tous, apporte une solution. Il suffit que l'utilisateur charge dans RPG Explorer les couches RPG relatives à son ou ses départements de travail⁵, la correspondance entre îlot et séquence se faisant par le numéro d'îlot. Une difficulté demeure toutefois. En effet, bien que les découpages d'îlots soient rigoureusement les mêmes sur le site de l'IGN ou dans les données issues de l'ASP, en revanche les numéros d'îlots utilisés diffèrent entre les deux sources de données. En amont de la transmission des données de séquence il faudra donc que l'ODR convertisse les numéros d'îlots ASP en numéros d'îlots IGN.

3.2. Choix de la période sur laquelle sont établies les séquences de groupes cultures

Sans autre développement informatique à l'ODR, on ne peut actuellement proposer au téléchargement que les séquences portant sur la période 2006-2014. Cela implique donc que l'utilisateur télécharge sur le site de l'IGN les données RPG de l'année 2014 pour disposer du shape des îlots permettant de localiser les séquences de l'ODR. Toute autre année de référence⁶ que 2014 ou autre année de démarrage de séquence que 2006 est actuellement impossible. Il est clair qu'une fois le module de téléchargement des données 2006-2014 mis en place il faudra rapidement implanter à l'ODR de nouveaux algorithmes de reconstitution des séquences (plus ou moins paramétrable, inspiré de RPG-E) permettant les traitements a minima sur les années postérieures à 2014. L'idéal étant que l'utilisateur ait toute latitude pour le choix des années sur lesquelles il souhaite calculer les séquences. On pourrait effectivement penser qu'il est toujours souhaitable de travailler à partir de l'année la plus récente sur l'ensemble de la profondeur temporelle disponible à ce jour, donc en remontant à 2006. Ce raisonnement se tient si on s'intéresse à la situation 2014 mais il pose problème si on s'intéresse à l'évolution des successions de culture ou des assolements de rotation dans le temps. Sur la période 2006-2014 il y a un nombre d'îlots conséquent pour lesquels les séquences ne sont pas connues sur l'ensemble de la période mais uniquement sur les 2 ou 3 premières années à compter de l'année de référence (ex 2014-2012). Si on souhaite faire une analyse de l'évolution des successions de cultures sur 2006-2014, il est donc préférable d'analyser 2006-2010 d'une part et 2010-2014 d'autre part. Travailler directement sur 2006-2014 pourrait conduire à des séquences de cultures biaisées sur la période la plus ancienne avec notamment beaucoup plus de séquences de prairies permanentes que la réalité. Toutefois, travailler sur 2006-2010 d'un côté et 2010-2014 de l'autre nécessite que l'utilisateur dispose du fichier de shape des îlots non seulement pour 2014 mais aussi pour 2010. Or le fichier shape 2010 n'est pas mis à disposition libre par l'IGN et en

⁴ Ce sont les mêmes données que celles de niveau 4 utilisées pour RPG Explorer mais sans aucune information sur les exploitations agricoles.

⁵ Rappelons que ce chargement de données se fait par région et que c'est RPG Explorer qui effectue la ventilation dans chacun des départements correspondants.

⁶ On appelle année de référence l'année la plus récente de la succession

l'absence d'études juridiques plus approfondie, il ne nous semble pas souhaitable que l'ODR diffuse un fichier shape acquis auprès de l'ASP dans le cadre d'une convention spécifique. Une solution possible à ce stade serait d'associer chacune des séquences non pas à un numéro d'îlot mais au point barycentre de cet îlot, les coordonnées spatiales de ce barycentre étant transmises avec la tables des séquences à l'utilisateur. La possibilité d'effectuer un choix sur l'année de démarrage et l'année de fin de la séquence serait une réelle plus-value mais ceci pose en retour une question importante au niveau de l'ODR à savoir quand et comment générer toute la gamme des séquences de groupes cultures que pourraient demander les utilisateurs.

3.3. Entre traitement anticipé ou traitement au fil de l'eau à l'ODR...

Dans le fonctionnement actuel, les séquences ont été calculées sur l'ensemble de la France pour une période donnée (2006-2014 par exemple) par l'UMR AGIR, puis le résultat a été stocké par l'US ODR et des extractions sont faites en fonction des demandes. Si comme proposé ci-dessus on laisse la possibilité à l'utilisateur de choisir les périodes sur lesquelles sont reconstituées des séquences, cela complique fortement la situation. Soit l'ODR effectuée au niveau de la France entière la reconstitution des séquences pour l'ensemble des combinaisons possibles. Rappelons sur que la période 2006-2014 cela correspond à 36 déterminations nationales de séquences de groupes cultures (8 possibilités avec 2014 comme année de référence, 7 avec 2013,...et 1 avec 2007).

Le traitement en amont pose aussi la question de l'actualisation des données de séquences. En effet les algorithmes de reconstitution des séquences tendent à s'améliorer au fil des réflexions des chercheurs de l'ODR et du SADAPT notamment. Par exemple actuellement (en 2018) une réflexion est en cours pour utiliser les découpages parcellaires fournis dans les données récentes (à partir de 2015) comme aide à la reconstruction des séquences pour les années antérieures à 2015. Cette prise en compte s'avère prometteuse mais en toute logique, si elle est implantée dans les nouvelles versions des outils de l'ODR et du SADAPT il faudrait que cela se retrouve aussi dans les séquences fournies aux utilisateurs finaux d'où autant de recalculs à l'échelon national. On peut envisager un autre système dans lequel les données sollicitées par les utilisateurs finaux seraient établies systématiquement avec l'algorithme le plus récent d'établissement des séquences. On pourrait ainsi avoir un système qui vérifie pour chacun des départements pour lesquels on a des données de séquences calculées si la période est déjà disponible et si l'algorithme avec lequel le calcul a été fait est bien le plus récent. Si c'est le cas les données pourraient être livrées telles quelles à l'utilisateur final sinon elles seraient calculées spécifiquement à cette occasion, stockées dans la base de l'ODR (éventuellement en remplacement d'anciennes données établies avec des algorithmes plus anciens) et mis à disposition des utilisateurs suivants. Cette solution bien que semblant plus prometteuse en théorie demandera à être expertisée plus finement.

3.4. Entre envoi en direct et envoi différé des données de séquence

Nous avons vu dans le point précédent que, face à une demande effectuée par un ordinateur équipé de RPG Explorer, on pouvait soit disposer des données directement soit avoir besoin d'un délai pour les calculer avant de les transmettre. Une solution possible serait que lors de la formulation de la demande en ligne, un « accusé » de réception de la base de l'ODR lui indique que sa demande a bien été reçue tout en lui spécifiant le statut des données demandées au moment où la demande a été reçue (ex : données disponibles ou données à calculer). RPG Explorer pourrait une fois ce message reçu interroger de lui-même la base selon un pas de temps régulier à définir. Une fois les données rendues disponibles elles seraient téléchargées par RPG Explorer qui pourrait les identifier directement et les intégrer dans les tables de la base locale. Une solution plus souple et moins intégrée pourrait être que l'utilisateur transmette une adresse mail avec en même temps que sa demande au serveur de l'ODR et que les données lui soient transmises sous forme d'un lien de téléchargement lui permettant de récupérer des fichiers qui seront ensuite reconnus par RPG Explorer pour les intégrer dans la base locale.

4. Création d'un module spécifique dans RPG Explorer

Comme nous l'avons vu dans le point précédent il sera nécessaire de développer un module spécifique permettant à l'utilisateur de RPG explorer (1) d'exprimer de manière standardisée la nature

des données qu'il sollicite auprès de l'ODR et (2) d'intégrer dans sa base locale les données transmises par le serveur de l'ODR.

4.1. Demande standardisée

Il faut notamment spécifier l'étendue spatiale et la profondeur temporelle. On pourrait ainsi proposer une interface permettant de construire un lot de requêtes avant que celles-ci ne soient envoyées à l'ODR. On pourrait ainsi croiser un lot de départements avec une série d'année, enregistrer cette première demande puis en faire une autre qui pourrait par exemple correspondre à une autre série d'années pour le même lot de départements. Pour limiter la charge du réseau on pourrait aussi offrir la possibilité de récupérer les données de séquence non pas au département mais à la commune. L'idée étant que les données seraient de fait toutes calculées à l'échelle du département mais que l'extraction qui en serait faite pourrait se limiter à un lot de communes. Ce mode de fonctionnement demande qu'un même découpage communal soit disponible au niveau de RPG Explorer et de la base ODR (les découpages communaux pouvant varier plus facilement que les découpages départementaux). Pour éviter le problème des versions de RPG Explorer, on pourrait envisager que le module ne fonctionne que si l'utilisateur est connecté à internet et donc que les découpages en communes proposés par l'interface soient un miroir de ceux disponibles à l'ODR. Une fois l'ensemble des requêtes espace*temps construites, l'utilisateur aurait à indiquer l'adresse mail à laquelle il souhaite recevoir les données. La requête une fois envoyée au serveur de l'ODR resterait stockée en mémoire de RPG Explorer avec possibilité pour l'utilisateur de l'éditer à nouveau en changeant tout ou partie de son contenu (par ex. nouvelle adresse mail si erreur à ce niveau sur la demande précédente).

4.2. Intégration dans la base locale des données issues de l'ODR

Dans le fonctionnement actuel de RPG Explorer les tables de séquences sont générées dans le dossier choisi par l'utilisateur. Il serait donc souhaitable qu'il en soit de même et que l'utilisateur puisse intégrer les données livrées dans le dossier de son choix. Cela montre l'intérêt de passer par un lien vers un fichier téléchargeable que l'utilisateur place ensuite dans le bon dossier correspondant pour RPG-E. Pour que l'intégration se fasse selon le formalisme de RPG Explorer, il est essentiel que les données relatives aux découpages spatiaux (ex données RPG 2014) soient déjà intégrées à la base pour les départements et les années considérées. Une fois cette vérification faite le module pourra générer les tables de séquence dans le dossier choisi par l'utilisateur. Il s'agit bien d'une génération et pas uniquement d'une intégration car lors de la détermination des séquences, RPG Explorer génère deux tables : une table géométrique (revient à un fichier de forme avec pour données attributaires du texte qui explicite les séquences associées) et une table de données simple (liste des séquences par numéro d'îlot). Les données transmises par l'ODR peuvent être de 2 natures. Pour des séquences établies avec une année de référence disponible sur le site de l'IGN, les données de séquences transmises sont dans une table simple. Il faut donc qu'au moment de l'intégration des données dans le dossier, RPG Explorer leur associe la deuxième table (géométrique) créée à partir des données RPG dont le système aura vérifié l'existence au préalable. Si on travaille avec des données antérieures à 2014 les données fournies par l'ODR seront sous forme d'un fichier *shape* de points. Les points étant les barycentres des polygones utilisés pour l'année de référence pour laquelle l'utilisateur n'a théoriquement pas les données RPG. La procédure de génération portera en ce cas sur la table « simple » de séquences en prenant comme identifiant des séquences les coordonnées des barycentres des îlots, ceci afin d'avoir des identifiants uniques sur une même année quel que soit le département. Il faudra vérifier par la suite si le travail avec des données en mode point et non plus en mode polygone pose problème sur les modules valorisant les séquences dans RPG Explorer. Si cela pose problème on peut réassocier les données aux découpages parcellaires de l'année la plus proche de l'année de référence des séquences. Cette procédure peut générer un certain nombre d'erreurs. Par exemple si des îlots ont été partagés en deux l'affectation des données de séquences portées par un point unique (barycentre) se fera alors à un seul des îlots. On peut toutefois supposer que sur un territoire de grande dimension ces erreurs resteraient limitées.

5. Modification de la base de l'ODR pour une ouverture vers RPG Explorer

L'ODR héberge des données sur les séquences sous deux types de gestionnaires de base de données, en MySQL et en PostgreSQL, qui sont hébergés sur des serveurs différents. RPG-E étant basé sur une base de données de type PostgreSQL, c'est le serveur correspondant qui serait mobilisé. Il est possible de développer requêteur auquel sont envoyé les demandes (avec comme paramètres les années et départements concernés), qui déclenchent ensuite les actions nécessaires : validation de la demande par un administrateur ; création des requêtes de création du fichier d'extraction (voir de l'algorithme de reconstitution des séquences, si on est dans la configuration idéale où on peut générer des séquences à la demande pour des combinaisons espace*temps indisponibles) ; dépôt du fichier sur un espace dédié sécurisé et envoi du lien de téléchargement. A priori ceci est tout à fait compatible et réalisable avec la plateforme actuelle. Toutefois si on utilise le serveur ODR lié à Postgre, deux problèmes se posent. D'abord un problème de concurrence en termes de puissance de calcul et d'espace de stockage entre les travaux de l'ODR et les besoins liées aux séquences de culture et à leur mobilisation par les utilisateurs de RPG-E. En effet il peut y avoir des phénomènes de congestion pénalisant les 2 aspects. D'autre part, la version de PostgreSQL utilisé par RPG-E et celle installée sur le serveur ODR ne sont pas forcément les mêmes, et leur évolution ne répond pas aux mêmes contraintes techniques et temporelles. Il paraît donc souhaitable de dédier un serveur spécifique pour la collaboration entre les équipes, ce qui implique donc un coût de location annuel pour le serveur et la disponibilité de compétences nécessaires pour administrer le serveur.

Une autre piste de rapprochement RPG-E/ODR serait de faire tourner le module de RPG-E qui génère les séquences directement via les serveurs et machines de l'ODR. Cette option est fort intéressante pour l'ensemble des partenaires, car elle est la seule voie réaliste à termes pour pérenniser les travaux. La pérennité de T-AGIR étant déjà mis à mal, se concentrer sur RPG-E, plus multi-usages et capable de gérer les données post-2014 du RPG, apparaît opportun, avant que le développement/maintenance du logiciel s'arrête faute de moyens humains et/ou financiers. Cela impliquerait de :

- Pouvoir extraire le module/composant de RPG Explorer qui est en charge de l'exécution du calcul de séquences
- Que ce module soit portable dans l'environnement du serveur que l'ODR va mettre à disposition (Linux Debian – mais dans lequel on peut aussi installer une machine virtuel windows, etc. Tout étant question de temps/moyens)
- Que l'exécution de ce module soit déclenchable par l'ODR (à la réception d'une nouvelle année RPG), si possible de manière automatique.

Pour l'instant, RPG-E établit un ensemble de requêtes PostgreSQL qui sont lancées ensuite sur un serveur local. Ces requêtes sont portables sur d'autres plateformes avec une même version de PostgreSQL. L'idée serait d'extraire de RPG-Explorer le module/composant qui exécute la tâche et le rendre fonctionnel en exécution de commandes (lignes de commande informatique) pour un serveur Linux Debian comme ceux dont dispose l'ODR. Mais le problème de portabilité vient des objets métiers générés en VB.net dans l'outil RPG-E et sur lesquels s'appuient les processus. Ces objets facilitent le travail de développement, mais réduisent la portabilité de l'outil, en particulier sur des serveurs Linux. Trois options de contournement sont envisageables :

- Option 1 : traduire les objets-métier en requêtes SQL
- Option 2 : passer de VB.net à un autre langage objet qui soient portables sur Linux (java, php, autres). Cela semble plus pertinent de garder la couche « objet » plutôt que de tout faire en requêtes SQL, en termes de « lourdeur » du code et de praticité pour les développements futurs.
- Option 3 : passer par un serveur pour les étapes d'intégration des données et de filiation (étapes qui se font via des requêtes SQL) ; puis connecter RPG-E au serveur ODR : le calcul se fera plus rapidement que d'habitude (sachant qu'en plus avec la parallélisation des requêtes, on peut optimiser l'utilisation concomitante de plusieurs processeurs). On passera par l'interface de RPG-E pour la reconnaissance des séquences, sur un ordi « customisé », avec renvoi dans le serveur les données en sortie, et qui pourront être ensuite distribuées aux différents utilisateurs.

C'est cette option 3 qui paraît la plus réaliste, en termes techniques et en termes de ressources disponibles. Elle implique toutefois que, dans un premier temps en tout cas, les versions de séquences disponibles dépendent des paramètres (année, durée, etc.) avec lesquels ils ont été produits au préalable par l'ODR ou autre partenaire ayant accès à l'interface de RPG-E relié au serveur. Le serveur étant dédiée à ces travaux et données « Séquences », il est toutefois

envisageable de faire tourner un ensemble de versions sélectionnées par des experts et couvrant la majorité des besoins.

Pour réaliser le présent cahier des charges, même dans son ambition « minimaliste », il faut des moyens humains significatifs et dédiés ; c'est-à-dire pas juste pour un « coup de main », et si possible une personne bénéficiant d'un statut pérenne pour réduire le risque de pertes de compétences nécessaires à l'utilisation, au maintien et au développement des outils informatiques. Pour approfondir le diagnostic et les pistes élaborées, traduire le cahier des charges en opérations techniques et les réaliser. Une des difficultés est que tant que le travail n'a pas commencé, on a du mal à mesurer l'envergure du chantier. De ce point de vue, des rapproches avec des équipes ayant des compétences poussées en Maths-Infos seraient un plus, notamment du côté du département INRA Mathématiques et Informatiques Appliquées.

6. Glossaire

Agence de Services et de Paiement (ASP) : l'ASP est chargé de verser les aides de la politique agricole commune (PAC) aux agriculteurs. A ce titre, elle gère un système d'information géographique informatisé (avec depuis 2015 l'appui de l'IGN) pour l'identification des îlots et la description de leur occupation culturale

Groupes de cultures (gc) : regroupement généré par l'ASP d'après les cultures déclarées sur l'îlot et les compositions des groupes de la nomenclature utilisée. La version disponible pour les travaux de recherche comprend 28 groupes de cultures différents.

Ilot: ensemble de parcelles culturales contiguës, entières ou partielles, portant une ou plusieurs cultures, exploitées par vous-même, limité par des éléments facilement repérables et permanents, comme un chemin, une route, un ruisseau...ou par d'autres exploitations. L'ensemble de ces parcelles sont stables d'une année sur l'autre. Ils constituent l'unité géographique de référence du RPG.

Recensement parcellaire graphique (RPG) : Le RPG est constitué de l'ensemble des parcelles culturales géolocalisées par les déclarants eux-mêmes (sur fond ortho-photographique fourni) lors de la campagne annuelle de "déclarations PAC" lancée au printemps sur l'ensemble du territoire français. Leur tracé sert de base légale à l'Agence de Services et de Paiement (ASP) pour déterminer le montant de diverses aides directes aux agriculteurs

RPG-E/ RPG-Explorer : procédure développée par l'UMR SADAPT pour le traitement des données du RPG et permettant notamment de reconstituer les séquences de culture dans les îlots RPG. RPG-Explorer inclut également d'autres fonctionnalités pour traiter et analyser les données du RPG (trajectoire d'exploitation, assolement de rotations, etc.)

Séquences de cultures : succession de cultures sur plusieurs années dans un îlot donné.

T-AGIR : procédure développée par l'UMR AGIR pour le traitement des données du RPG et permettant de reconstituer les séquences de culture dans les îlots RPG

7. Sigles & Abréviations

ASP : Agence de Services et de Paiement

IGN : Institut Géographique National

Inra : Institut National de la Recherche Agronomique

RPG : recensement parcellaire graphique

UMR AGIR : Unité mixte de recherche Agronomie, Innovations et Territoires

UMR SADAPT : Unité mixte de recherche Sciences Action Développement – Activités Produits
Territoires

US ODR : Unité de service Observatoire du Développement Rural

Avec le soutien financier de

**AGENCE FRANÇAISE
POUR LA BIODIVERSITÉ**
ÉTABLISSEMENT PUBLIC DE L'ÉTAT



www.agence-francaise-biodiversite.fr