



HAL
open science

Adaptation aux paysages agricoles européens d'une méthode cartographique d'analyse de la configuration et de la composition des cultures

François Calatayud

► **To cite this version:**

François Calatayud. Adaptation aux paysages agricoles européens d'une méthode cartographique d'analyse de la configuration et de la composition des cultures. Sciences du Vivant [q-bio]. 2011. hal-02810791

HAL Id: hal-02810791

<https://hal.inrae.fr/hal-02810791>

Submitted on 6 Jun 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.



Université de Toulouse

MASTER 2 GEOMATIQUE

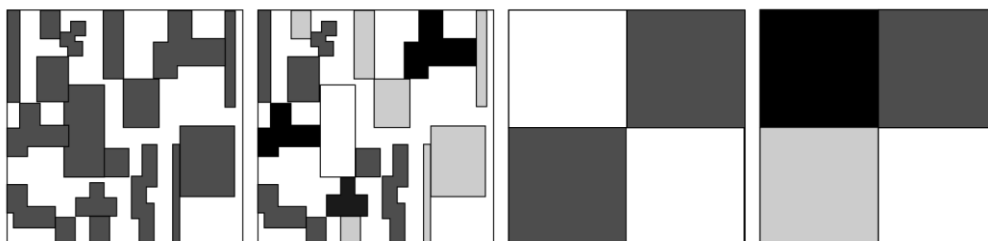
Parcours Professionnel :

« **S**cience de l'**I**nformation **G**éoréférencée pour la **M**aîtrise de
l'environnement et l'**A**ménagement des territoires » (**SIGMA**)

<http://sigma.univ-toulouse.fr>

RAPPORT DE STAGE

Adaptation aux paysages agricoles européens
d'une méthode cartographique d'analyse de la
configuration et de la composition des cultures



Calatayud François

Unité mixte de Recherche DYNAFOR / INRA / INP-ENSAT



Maître de stage : Sylvie Ladet (encadrant principal) et Gérard Balent
Tuteur-enseignant : David Sheeren

Septembre 2011

Résumé :

Les paysages agricoles occupent près de 40% de la superficie des terres disponibles. A ce titre, ils jouent un rôle important dans la conservation de la biodiversité. En retour, cette dernière contribue à la production agricole en fournissant des services écosystémiques tels que la pollinisation ou la lutte biologique. Le rôle de l'hétérogénéité du paysage agricole dans le maintien de la biodiversité se révèle être une piste de recherche intéressante qui pourrait suggérer de nouvelles politiques agricoles visant à favoriser la biodiversité. Le projet FARMLAND (BiodivERSa2011-66) s'insère dans ce contexte en s'inspirant de travaux récents combinant des outils de géomatique, de télédétection et de géostatistique. Sur la base du calcul d'indices spatiaux, ces travaux proposent une méthode pour réaliser la sélection de quadrats permettant de considérer l'influence sur la biodiversité, de l'hétérogénéité de composition du paysage (nombre et proportions des différents types de couvert) et de l'hétérogénéité de configuration (arrangement spatial des types de couvert). Ce projet doit aboutir à la sélection d'une cinquantaine de carrés dans 7 sites européens au sein desquels seront effectués des relevés de biodiversité et des évaluations des services écosystémiques. Nous présentons les différentes étapes de la méthode cartographique (classification et segmentation d'image satellitaire, analyse par fenêtre mobile, requête multicritère) menant à la réalisation du plan d'échantillonnage pour le site des Vallées et Coteaux de Gascogne qui sert de zone test pour les autres sites européens. Nous discutons des adaptations nécessaires et des différences avec le site de référence canadien.

Abstract :

Agricultural landscapes occupy 40% of the available land area, and as such play an important role in conserving biodiversity. Farmland biodiversity makes an important contribution to agricultural production by providing ecosystem services such as crop pollination and biological control of crop pests. The role of agricultural landscape heterogeneity in maintaining biodiversity appears to be a promising research field that may suggest new agricultural policies. Based on recent works combining geomatic, remote sensing and geostatistics, Farmland project (BiodivERSa2011-66) propose a new methodology for sampling quadrats that would allow to consider the impact of compositional heterogeneity (number and proportion of cover types) and configurational heterogeneity (spatial disposition) on biodiversity and ecosystem services. About fifty quadrats will be sampled in each of the 7 participating European countries in which biodiversity survey and ecosystem services evaluations will be undertaken. We present the different steps of the methodology (including classifications, feature extraction, moving window and multi-criteria analyses) that lead to the sampling procedure for the LTER site of Gascony (Southwestern France). We discuss discrepancies with the Canadian reference site and suggest adaptations to European landscape characteristics.

Remerciements

Ce travail s'inscrit dans un contexte particulier puisqu'il s'est déroulé dans le cadre d'un changement de vie à plus d'un titre.

Il y a un peu plus d'une année, je décidais de changer de métier en mettant entre parenthèse mon activité de chargé de recherches sur le comportement animal. Ayant toujours été attiré par l'analyse de données, je choisisais de m'orienter vers la géomatique, discipline clef en écologie depuis plusieurs années. Après avoir abordé la question avec ma hiérarchie au sein du département Ecologie des Forêts, Prairies et Milieux Aquatiques (EFPA), j'ai contacté Daphné Frullini, responsable Formation Permanente du centre INRA de Toulouse. Elle m'a orienté vers les chercheurs et ingénieurs de L'UMR DYNAFOR et plus particulièrement vers Sylvie Ladet (Ingénieure d'étude) et Gérard Balent (alors Directeur de l'unité). Le contact est particulièrement bien passé et la possibilité de m'accueillir dans le laboratoire s'est présentée. Il a alors fallu monter un dossier de demande de Formation Diplomante en un temps record avec l'accord et le soutien de la direction du département EFPA. Ces échanges ont plus particulièrement impliqué le chef de département Jean-Marc Guehl et Michel Verger. Sur les conseils de Sylvie Ladet, j'ai alors proposé ma candidature pour suivre la formation dispensée par l'équipe pédagogique du Master2 SIGMA. Cette demande a été accueillie favorablement ainsi que ma demande de Formation Diplomante et l'aventure commençait...

Cette aventure n'a pas été de tout repos, mais s'est révélée passionnante. J'ai pu goûter à nouveau au plaisir des salles de cours de la fac avec l'impression de les avoir quittées la veille et redécouvrir l'ambiance d'une promo fraîchement constituée. Cette ambiance n'a pas été le moindre des atouts de cette année car sans elle et sans le lien qui nous a uni, les différentes épreuves que nous avons traversées auraient été autrement plus rudes.

Puis une page de plus s'est inscrite dans ma vie et dans celle de Sandrine, ma compagne, avec l'arrivée de notre petite Félipa. Très tôt cette dernière a eu la présence d'esprit de contribuer, avec Sandrine, au bon déroulement de ma formation. Sa naissance après les examens et ses nuits de plus de 10 heures d'affilées ont grandement aidé.

Puis le stage a commencé et j'ai découvert une nouvelle vie de laboratoire et le plaisir d'évoluer dans un contexte de recherche avec des interactions riches entre ses différents acteurs. Je pense particulièrement à Gérard Balent, à Sylvie Ladet et à David Sheeren, mes encadrants qui ont toujours été là pour répondre à mes questions, à Lenore Fahrig qui répondait si promptement (et en français !) à mes mails depuis le Canada, à Michel Goulard, mon voisin de couloir, qui lui aussi invective son PC ou le félicite selon l'humeur de celui-ci et qui a pris le temps de répondre à mes questions sur les statistiques, à Valéry Rasplus qui a relu mon mémoire. Une pensée toute particulière va à Sandrine qui, bien qu'extérieure au laboratoire, a mis ses compétences en analyse de données et en programmation à mon service.

Que toutes les personnes citées ici soient sincèrement remerciées (et que celles que j'ai omises, par distraction ou faute de place ne m'en veuillent pas trop...).

Sommaire

I. Introduction.....	1
II. Contexte de l'étude.....	2
1) Une nouvelle perspective sur la protection de la biodiversité : mettre l'agriculture au cœur du système.	2
2) Rôle de l'hétérogénéité du paysage.....	3
3) Vers une vision plurielle de l'hétérogénéité.	3
4) Des conséquences écologiques et politiques différentes.	5
5) Le projet FARMLAND.....	5
III. Objectifs.....	7
IV. Présentation de la chaîne de traitement et choix des données de départ.....	8
1) Aire d'étude.....	8
2) Choix des données de départ.....	8
3) Présentation générale de la chaîne de traitement.....	11
4) Plateforme logicielle.....	14
V. Etape 1 : Classification de l'occupation du sol pour définir la composition du paysage étudié.....	14
1) Utilisation de données agricoles du Registre Parcellaire Graphique (RPG) pour réaliser des échantillons d'apprentissage et de test de la classification.....	14
2) Une étude préliminaire.	16
3) Discussion des nouveaux résultats de la classification.	16
VI. Etape 2 : Segmentation vectorielle des limites de parcelles pour définir la configuration du paysage étudié.	19
1) Introduction aux méthodes de segmentation.....	19
2) Résultats et validation des segmentations.....	19
3) Discussion des résultats de la segmentation.	22
VII. Etape 3 : Prétraitement pour l'analyse par fenêtre mobile avec FRAGSTATS v3.3.	23
1) Présentation du logiciel FRAGSTATS v3.3.....	23
2) Préparation aux traitements par fenêtre mobile.....	23
VIII. Etape 4 : Analyses par fenêtre mobile.	25
1) Détermination des différentes échelles d'analyse.....	25
2) Calcul des masques retenant entre 60 et 90 % de surface agricole.....	26
3) Calcul des rasters basés sur la taille moyenne de <i>patch</i>	26
4) Calcul des rasters basés sur l'indice de diversité de Shannon.	27

5) Estimation des rasters à conserver : choix de la taille de la fenêtre mobile.....	28
6) Estimation de la distance minimale entre deux fenêtres mobiles.....	34
7) Discussion sur le choix des tailles de fenêtre mobile à conserver et sur l'estimation de la distance minimale entre centre des quadrats d'échantillonnage.....	36
IX. Etape 5 : Echantillonnage des points de relevés des suivis faunistiques et floristiques.....	36
1) Détermination des plages de valeur correspondant aux quatre combinaisons de modalités d'hétérogénéité et recodage des rasters d'hétérogénéité.....	37
2) Echantillonnage des barycentres des quadrats.....	39
3) Vérification terrain.....	41
X. Conclusions générales et perspectives.....	47
XI. Références bibliographiques.....	50
ANNEXES.....	53
FICHE QUALITE UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	54
PRESENTATION DES METRIQUES INDICE DE DIVERSITE DE SHANNON ET TAILLE MOYENNE DE TACHE	77
PRESENTATION DE L'INDICE D'AUTOCORRELATION DE MORAN.....	79

I. Introduction.

Ce travail s'inscrit dans le cadre du projet de recherche FARMLAND (BiodivERsA2011-66) visant à mettre en relation l'hétérogénéité de paysages agricoles européens et la biodiversité. De nombreuses études récentes ont en effet mis en avant cette piste de recherche et suggèrent des modifications des pratiques agricoles qui soient acceptées par les acteurs de l'espace rural. Cet aspect sociologique serait une des conditions essentielles de l'efficacité de nouvelles politiques agricoles permettant la préservation et la promotion de la biodiversité.

Dans cette perspective, ce projet européen s'est largement inspiré des travaux récents d'une équipe Canadienne de recherche en écologie du paysage sous la direction de Lenore Fahrig (qui contribue également à FARMLAND). Plus particulièrement, ce travail se construit sur la base d'un article (Pasher *et al.*, draft) soumis à la revue *Landscape Ecology* à la fin de l'été 2011. Cet article pose les différentes étapes d'une chaîne de traitement amenant à la sélection de lieux sur lesquels effectuer des relevés de biodiversité selon différents critères d'occupation du sol. Cette chaîne de traitement nécessite alors l'utilisation de tout un panel d'outils méthodologiques de la géomatique, de la télédétection et de géostatistiques afin de déterminer ces critères. Cela implique notamment d'adapter cette chaîne de traitement à la situation particulière de chaque site d'étude.

II. Contexte de l'étude.

1) Une nouvelle perspective sur la protection de la biodiversité : mettre l'agriculture au cœur du système.

Les études pointant une diminution drastique de la biodiversité ou annonçant celle-ci à brève échéance sont désormais fréquemment mises en avant (Smart *et al.*, 2006 ; Davidson *et al.*, 2009). Une recherche bibliographique avec les mots clefs « *biodiversity loss* » sur la plateforme *ScienceDirect* propose plus de 28000 réponses sur les dix dernières années dont l'évolution par année est retracée dans la figure 1 :

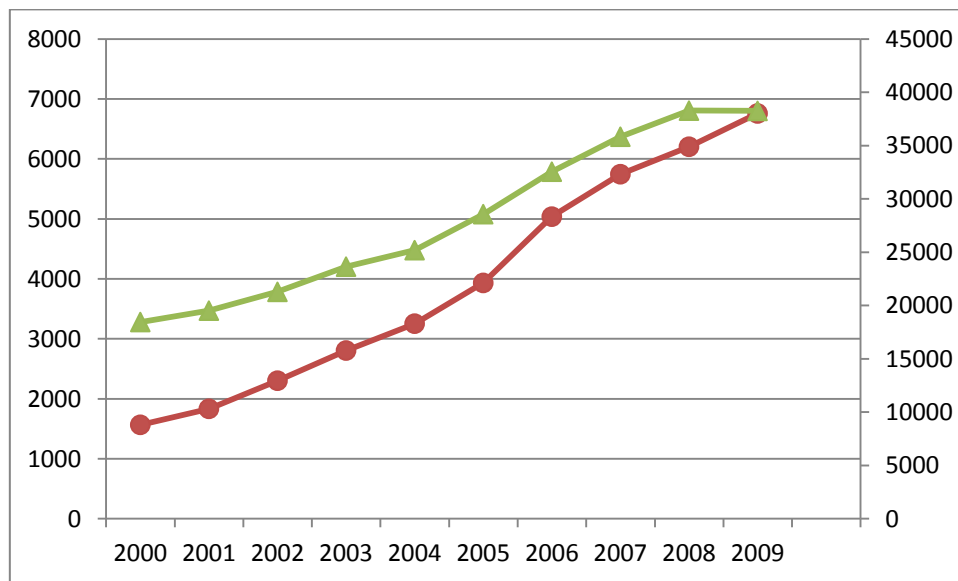


Figure 1: Evolution du nombre de titres d'articles ou de chapitres d'ouvrages comportant les termes « *biodiversity loss* » (marqueurs ronds en rouge et axe des ordonnées de gauche) dans un des champs indexés dans la base de la plateforme bibliographique *ScienceDirect* et du nombre de titres d'articles ou de chapitres d'ouvrages comportant le terme « *ecology* » (marqueurs triangulaires en vert et axe des ordonnées de droite).

Entre la première et la dernière année de cette période, le nombre de titres connaît une augmentation d'un facteur de 331% alors que comparativement le nombre de titres comportant le terme « *ecology* » peine à doubler. Ainsi, même en tenant compte de l'augmentation globale de la production scientifique en écologie, il apparaît que la thématique de la diminution de la biodiversité est particulièrement importante et d'actualité.

Bien que certains scientifiques, hommes politiques ou plus largement acteurs de l'espace public contestent l'origine humaine de cette perte de biodiversité, un lien direct existe entre l'activité anthropique et un impact sur l'environnement (Smith *et al.*, 2008 ; Srinivasan *et al.*, 2008). Cela a d'ailleurs amené des chercheurs à proposer une nouvelle époque, l'anthropocène, à l'échelle des temps géologiques justement basé sur l'empreinte écologique de l'espèce humaine (Zalasiewicz *et al.*, 2011). Ce lien est particulièrement net dans le cadre des pratiques agricoles qui interviennent sur approximativement 40% des surfaces de terre disponibles (Foley *et al.*, 2005 ; Smith *et al.*, 2008 ; Geiger *et al.*, 2010). Toutefois, différentes études en écologie du paysage, bien qu'allant dans ce sens, proposent de tirer parti de ce lien plutôt que de se borner à une mise en accusation (Benton, Vickery et

Wilson 2003 ; Tschardtke *et al.*, 2005 ; Fahrig *et al.*, 2010). En effet, si ce lien peut être mis en évidence, alors il pourrait être possible de l'utiliser pour créer des conditions de gestion de la biodiversité promouvant son maintien voire son expansion (Altieri, 1999 ; Gurr, Wratten et Luna, 2003, Fahrig *et al.*, 2010). Cette biodiversité accrue pourrait alors se révéler particulièrement avantageuse pour les pratiques agricoles au travers de services écosystémiques en favorisant la pollinisation ou en limitant les déprédations occasionnées par les ravageurs des cultures (Gurr, Wratten et Luna, 2003).

2) Rôle de l'hétérogénéité du paysage.

Différentes politiques d'interventions impliquant une modification des pratiques agricoles ont été envisagées. A côté de solutions relativement classiques comme la diminution des intrants ou l'augmentation de la part des éléments semi-naturels (bordures, haies, etc.) de nombreux auteurs proposent de prendre en compte l'hétérogénéité du paysage agricole dans le maintien et le développement de la biodiversité.

En effet, différentes études mettent en avant l'influence de l'hétérogénéité du paysage sur la biodiversité, que ce soit à l'aune de la dynamique des populations en place (Fraterrigo, Pearson et Turner, 2009) ou au regard des interactions entre espèces (Polis, Power et Huxel, 2004 cités par Fahrig *et al.*, 2010). Cette influence serait particulièrement importante dans le cas dans les écosystèmes agricoles (Benton, *et al.*, 2003 ; Tschardtke *et al.*, 2005 ; Le Roux *et al.*, 2008 ; Fahrig *et al.*, 2010). Gurr, Wratten et Luna (2003) mettent en écho deux citations : une phrase d'Odum (1971) questionnant (de manière peu poétique, il est vrai) le rôle des espèces « qui ne se mangent ou ne se vendent pas », une autre de Thies et Tschardtke (1999) qui fait le pendant de la première en lui apportant une réponse : « *In structurally complex landscapes, parasitism was higher and crop damage was lower than in simple landscapes...* ». L'hétérogénéité du paysage pourrait donc recéler une des clefs au problème de l'impact de l'agriculture sur la biodiversité via la production de services écosystémiques. On pourrait alors imaginer que l'utilisation de l'hétérogénéité du paysage pour créer des conditions favorables à la biodiversité crée également des conditions favorables à la limitation de l'usage d'une partie des intrants, aboutissant à l'instauration d'un cycle vertueux. Il apparaît donc crucial de mieux connaître les relations entre biodiversité et hétérogénéité du paysage.

3) Vers une vision plurielle de l'hétérogénéité.

Dans plusieurs travaux récents Fahrig a introduit, défini et précisé ce qu'était l'hétérogénéité du paysage (Fahrig et Nuttle, 2005 ; Fahrig *et al.* 2010 et Pasher *et al.*, draft). L'hétérogénéité de composition est relative au changement quantitatif du nombre d'espèces présentes dans le paysage (et non plus seulement à l'échelle de l'habitat). L'hétérogénéité de configuration est relative à la complexité de forme et de distribution spatiale des différents habitats composant le paysage. Dans son paradigme, cette approche s'inspire directement des travaux de Duelli (1992, 1997) et plus particulièrement de son concept de mosaïque. Selon cet auteur, les facteurs les plus pertinents pour prédire et évaluer la biodiversité portent sur la variabilité des habitats (comme le nombre d'habitats différents pour une aire d'étude donnée) et sur l'hétérogénéité des habitats (intégrant la complexité de forme des habitats).

D'après ces auteurs, l'hétérogénéité du paysage se conjugue au pluriel. Les études basées sur les relations entre la biodiversité et l'hétérogénéité doivent en conséquence non seulement prendre en compte les effets premiers de chacune des composantes de l'hétérogénéité, mais également leurs interactions. La figure 2 inspirée de Duelli (1997) et Fahrig *et al.*, 2010 explicite les différents cas de figure à prendre en compte.

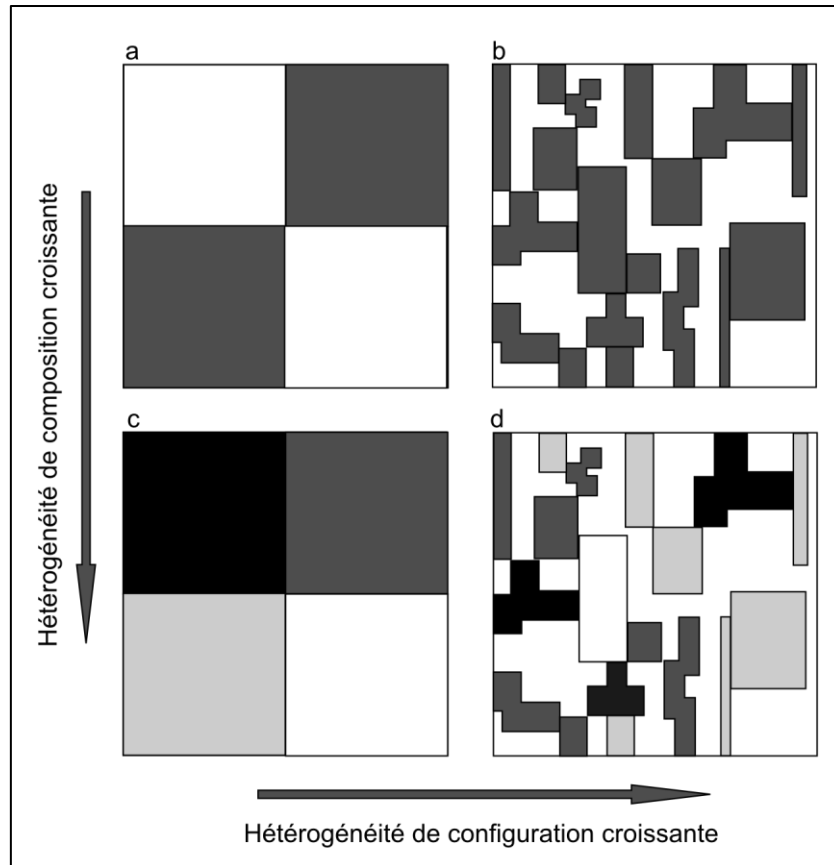


Figure 2: Représentation des différentes combinaisons des modalités d'hétérogénéité de composition (nombre de carré de couleurs différentes) et de configuration (surface moyenne des carrés) retenues par la suite (a : Faible/Faible ; b : Fort/Faible ; c : Faible/Fort ; d : Fort/Fort). (d'après Fahrig *et al.*, 2010).

Fahrig *et al.* (2010) proposent d'étudier les relations entre hétérogénéité du paysage et biodiversité, d'une part dans les milieux agricoles et d'autre part dans les milieux « plus naturels », le degré de naturalité se distinguant par différents critères :

- La part de la production primaire consommée par des humains ou leur bétail.
- La présence d'espèces ayant une histoire co-évolutive ou pour le moins une association à long terme.
- La fréquence et l'intensité des perturbations anthropiques.

Afin d'établir une limite dans le continuum observable entre ces deux pôles, Fahrig *et al.* (2010) suggère de poser la frontière de « la plus grande naturalité » à 60% d'occupation du sol en production agricole selon ces critères ce qui correspond aux cultures mais aussi aux prairies temporaires et permanentes.

4) Des conséquences écologiques et politiques différentes.

L'effet de l'hétérogénéité d'un paysage semi-naturel sur sa biodiversité semble avoir surtout été envisagé par le biais de l'hétérogénéité de composition à savoir la présence/absence ou la quantité d'un certain type d'occupation du sol (Ouin et Burel, 2002 ; Tschardt *et al.*, 2005). Dans ce cas, cette augmentation de biodiversité peut notamment trouver une explication dans l'accroissement des interactions entre espèces, et la persistance de méta-communautés (Fahrig *et al.*, 2010). Un autre effet de cette hétérogénéité pourrait être lié au fait que certaines espèces ont des préférences ou des besoins différents en fonction de la période de leur cycle de vie selon la théorie de la « complémentation du paysage » (Dunning *et al.*, 1992). Les effets de l'hétérogénéité de configuration devraient, eux aussi, utiliser le principe de la complémentation du paysage de Dunning. Toutefois, les deux dimensions de l'hétérogénéité pourraient se distinguer sur certains points. Notamment, une trop grande hétérogénéité de composition pourrait amener à une diminution des surfaces de chaque modalité d'occupation du sol et à une diminution de la présence des espèces spécialisées à certaines d'entre elles. Les principes à l'origine des effets de l'hétérogénéité des paysages semi-naturels sur leur biodiversité devraient également être sous-jacents à la dynamique des paysages agricoles (persistance des méta-communautés, complémentation du paysage), et ce quelle que soit la dimension de l'hétérogénéité prise en compte.

Toutefois, les conséquences, en termes de coûts, des politiques agricoles devraient être différentes selon la dimension mobilisée. Selon Fahrig *et al.* (2010) une action promouvant l'utilisation de l'hétérogénéité de composition (en augmentant le nombre de cultures différentes pour chaque agriculteur) pourrait avoir des conséquences en termes de modalités de stockage des différentes récoltes. Une action basée sur l'hétérogénéité de configuration (des parcelles de surface plus petite) devrait avoir des conséquences en termes de facilité de gestion des cultures (plus de visites avec des engins plus petits). Toutefois, au delà de l'aspect financier, la prise en compte de la dimension sociologique de ces pratiques semble être particulièrement importante dans la mesure où celle-ci conditionne directement le succès d'une politique de gestion de la biodiversité par les pratiques agricoles. Des actions intervenant sur l'hétérogénéité de composition et/ou de configuration du paysage pourraient justement s'inscrire dans ce capital culturel et social déjà existant et créer un lien entre l'activité professionnel des agriculteurs et la présence de certaines espèces clefs (Burton, Kuczera et Schwartz, 2008). Les chances de succès d'une politique de maintien et de promotion de la biodiversité seraient plus grandes en faisant de celle-ci un produit de la pratique agricole et un motif de prestige pour les acteurs de l'espace rural.

5) Le projet FARMLAND.

C'est précisément dans ce contexte que le projet Européen FARMLAND BiodivERsA2011-66 a été lancé, sur la base des travaux récents de Lenore Fahrig (Fahrig, 2003 ; Fahrig et Nuttle, 2005 ; Fahrig *et al.*, 2010 ; Pasher *et al.*, draft). Dans ses derniers articles elle propose un protocole permettant d'étudier les relations entre la biodiversité et les différentes dimensions de l'hétérogénéité du paysage de manière indépendante et en interaction. Ce protocole a permis à Pasher *et al.* de sélectionner 4 échantillons de 25

quadrats¹ correspondants aux quatre combinaisons de modalités d'hétérogénéité de composition et de configuration présentées précédemment dans la figure 2). Cette sélection de quadrats s'est faite au terme de toute une chaîne de traitement géomatique permettant de caractériser des paysages agricoles du Canada sur la base de leur occupation du sol. Le projet FARMLAND vise à transposer ce protocole aux paysages européens en tenant compte des spécificités locales pouvant être radicalement différentes de leurs homologues nord-américains. Afin de prendre en compte un panel suffisamment large de situations, différents sites d'étude ont été choisis sur un axe nord-sud de l'union européenne : un en Allemagne, quatre en France et un en Espagne (Figure 3).

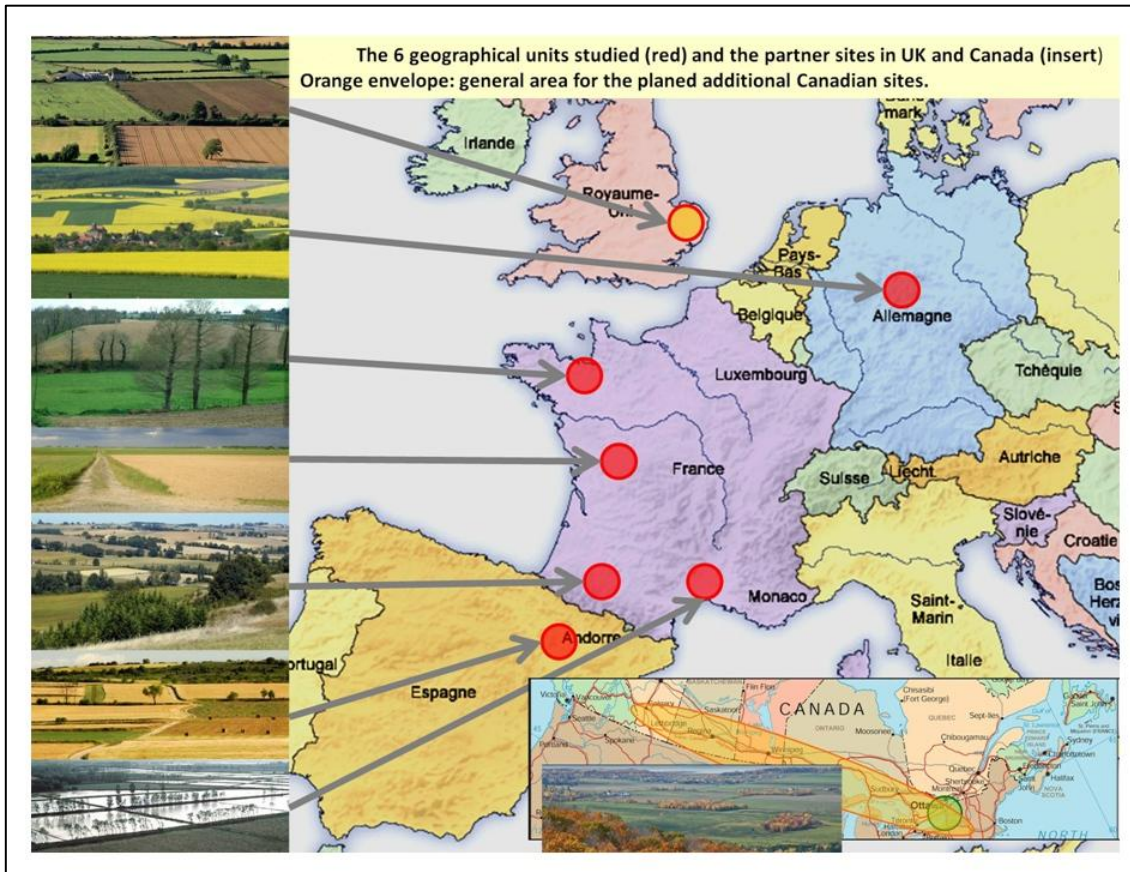


Figure 3: Répartition des sites d'études FARMLAND en Europe et localisation de l'étude princeps au Canada.

Ce projet regroupe huit équipes de recherche réparties dans différents pays européens et l'équipe Canadienne initiatrice du projet:

- Allemagne :
 - Université de Göttingen (sous la direction de Teja Tschardt)
- Royaume-Uni :
 - British Trust for Ornithology (sous la direction de Gavin Siriwardena)

¹ Méthode de recensement de la flore ou de la faune basée sur un quadrillage d'une zone à étudier. Ici les quadrats sont des carrés de taille constante positionnés selon un plan d'échantillonnage.

- France :
 - CEFE UMR 5175 (sous la direction de Jean-Louis Martin)
 - CEBC CNRS (sous la direction de Vincent Bretagnolle)
 - OSUR CNRS et Université de Rennes (sous la direction de Françoise Burel).
 - La Fondation Tour du Valat (sous la direction de Brigitte Poulin).
 - DYNAFOR UMR 1201 (sous la direction de Gérard Balent)
- Espagne :
- Canada :
 - Geomatics and Landscape Ecology Laboratory, Carleton University (sous la direction de Lenore Fahrig).

Ces équipes auront à répondre à quatre grandes questions dans le cadre de ce projet :

- Quels sont les changements de biodiversité qui accompagnent une augmentation de l'hétérogénéité de composition des paysages agricoles?
- Quels sont les changements de biodiversité qui accompagnent une augmentation de l'hétérogénéité de configuration des paysages agricoles ?
- Quelles mesures de la biodiversité sont les plus reliées aux services écosystémiques culturels (bénéfices non matériels telles la valeur culturelle et esthétique), de maintien, de régulation (pollinisation, contrôle des ravageurs de cultures) ?
- Quelle est la dimension sociologique des liens entre l'agriculture, l'hétérogénéité du paysage et la biodiversité ? Comment proposer des pratiques et des lignes de conduites acceptables qui créeront les structures spatiales permettant de promouvoir les services écosystémiques basés sur la biodiversité ?

La réponse à ces différentes questions comprendra donc une dimension écologique au travers de relevés de biodiversité impliquant différents taxons (plantes, coléoptères carabidés, arachnides, hyménoptères apoïdes, diptères syrphidés, lépidoptères rhopalocères, oiseaux et vers de terre), et une dimension sociologique. L'étape préliminaire de ce projet FARMLAND consistera à déterminer les différents quadrats sur lesquels effectuer ces relevés de biodiversité et correspondants aux différentes combinaisons de modalités d'hétérogénéité du paysage en adaptant le protocole initial aux paysages européens.

III. Objectifs.

Les objectifs de ce travail qui combine télédétection, analyse spatiale et géostatistique s'inscrivent autour de 3 axes:

- Réaliser une cartographie représentative de l'hétérogénéité de composition et de configuration d'un paysage agricole de large étendue suite à une chaîne de traitement.
- Déterminer des points d'échantillonnage pour de futures campagnes de relevés faunistiques et floristiques.

- Adopter un regard critique sur les différentes étapes par lesquelles nous sommes passés et effectuer des propositions afin de l'adapter aux caractéristiques des paysages que nous² aurons à étudier dans FARMLAND.

Dans la mesure où nous avons été amenés à parcourir une chaîne de traitement relativement longue que nous avons du adapter à notre situation, nous avons choisi de ne pas présenter ce travail selon un plan classique « introduction- matériel et méthode- résultat-discussion conclusion ». Il pourrait être plus intéressant et plus confortable pour le lecteur de présenter à grand trait la chaîne de traitement et de reprendre ce plan classique pour chaque étape.

IV. Présentation de la chaîne de traitement et choix des données de départ.

1) Aire d'étude.

L'aire d'étude sur laquelle est basé notre travail constitue l'un des quatre sites français du projet FARMLAND présentés dans la figure 3. L'emprise de ce site est localisée à cheval sur trois départements, l'Ariège, la Haute-Garonne et le Gers, bien que seuls les deux derniers (et principalement la Haute-Garonne) comprennent les sites échantillonnés en pratique. C'est en effet sur les terrasses de la rive gauche de la Garonne et les coteaux y adossés que les travaux de l'UMR DYNAFOR portent depuis plusieurs années et qu'un capital de données a été accumulé. On y retrouve notamment un observatoire LTER (*Long Term Ecological Research*) de suivi à long terme du paysage et de l'écologie ainsi qu'une zone ayant fait l'objet en 2010 d'une mission d'acquisition de données hyperspectrales et LIDAR par capteur aéroporté (MADONNA). Sur cette aire d'étude, les pratiques agricoles et pastorales ont façonné des paysages typiques constitués de grandes cultures, de bois et de prairies liées à l'élevage bovin (Choisis *et al.*, 2010). Les exploitations agricoles ont des systèmes de production très différents. Certaines se sont spécialisées en grandes cultures (blé, tournesol, maïs...), d'autres en élevage, en particulier dans la production de lait ou de viande bovine. D'autres encore ont gardé un système de production diversifié dit de polyculture-élevage, alliant au sein de l'exploitation un élevage bovin et des grandes cultures. La surface agricole utile moyenne est aujourd'hui comprise entre 20 et 180 hectares selon les systèmes de production (Choisis *et al.*, 2010).

2) Choix des données de départ.

Afin de réaliser cette étude basée sur l'occupation du sol dans le cadre des pratiques agricoles, nous devons avoir connaissance des différents types de cultures présentes sur l'aire d'étude.

- Le Registre Parcellaire Graphique (RPG) permet d'obtenir une certaine base de renseignement concernant les différentes cultures présentes sur notre aire d'étude. Ce document administratif de la Politique Agricole Commune est un système d'information administré par l'Agence de Services et de Paiement. Cette base de

² Je m'inscris dans le « nous » puisque pour rappel en tant qu'agent INRA titulaire, je reste membre de l'unité DYNAFOR au moins jusqu'en 2014 et je suis d'ores et déjà impliqué dans le projet FARMLAND

données annuelle va comprendre environ six millions d'îlots³ qui correspondent à des parcelles culturales contiguës exploitées par un même agriculteur. Celui-ci remplit chaque année un document renseignant la position des parcelles et l'occupation du sol (figure 4).

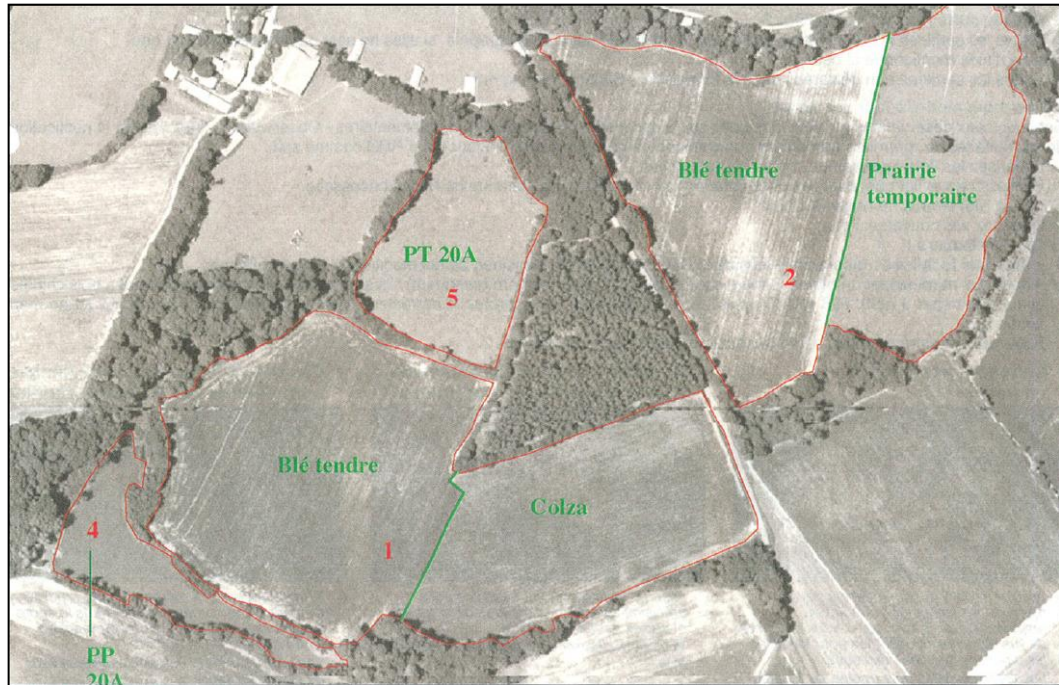


Figure 4 : Illustration du RPG sur quelques îlots en rouge avec les cultures renseignées en vert (source : Note d'utilisation du Registre Parcellaire Graphique 2005). Nota PT=prairie temporaire et PP= prairie permanente

Certains de ces îlots seront dit homogènes car constitués d'une seule modalité d'occupation du sol, d'autres seront dit hétérogènes car faisant intervenir différents types de cultures. Différents recodages de la liste des types de cultures présentes dans notre aire d'étude ont été nécessaires afin d'améliorer les performances de nos analyses et seront présentés ultérieurement.

- Le calendrier culturel permet de compléter les informations données par le Registre Parcellaire Graphique. Ce calendrier couvre une durée de 12 mois, et est décalé par rapport à l'année civile (d'octobre à septembre). Ce calendrier mettant en évidence le décalage temporel existant entre les cycles phénologiques des principales cultures retrouvées dans les coteaux de Gascogne, nous avons choisi de baser notre étude de l'occupation du sol sur plusieurs images satellites obtenues à différentes périodes réparties sur le calendrier culturel de manière à suivre les cycles de croissance de ces cultures prédominantes. La figure 5 permet d'associer ces cycles phénologiques à différentes prises de vues satellites.

³ La définition de l'îlot est donnée par la PAC. En vocabulaire de géomaticien, un îlot correspond à un polygone rassemblant des parcelles culturales exploitées par un même agriculteur sans obstacle de type route et ruisseau. Le dessin des îlots est mis à jour chaque année par les agriculteurs sur la base de l'orthophotographie de l'Institut Géographique National (BD ORTHO® de l'IGN).

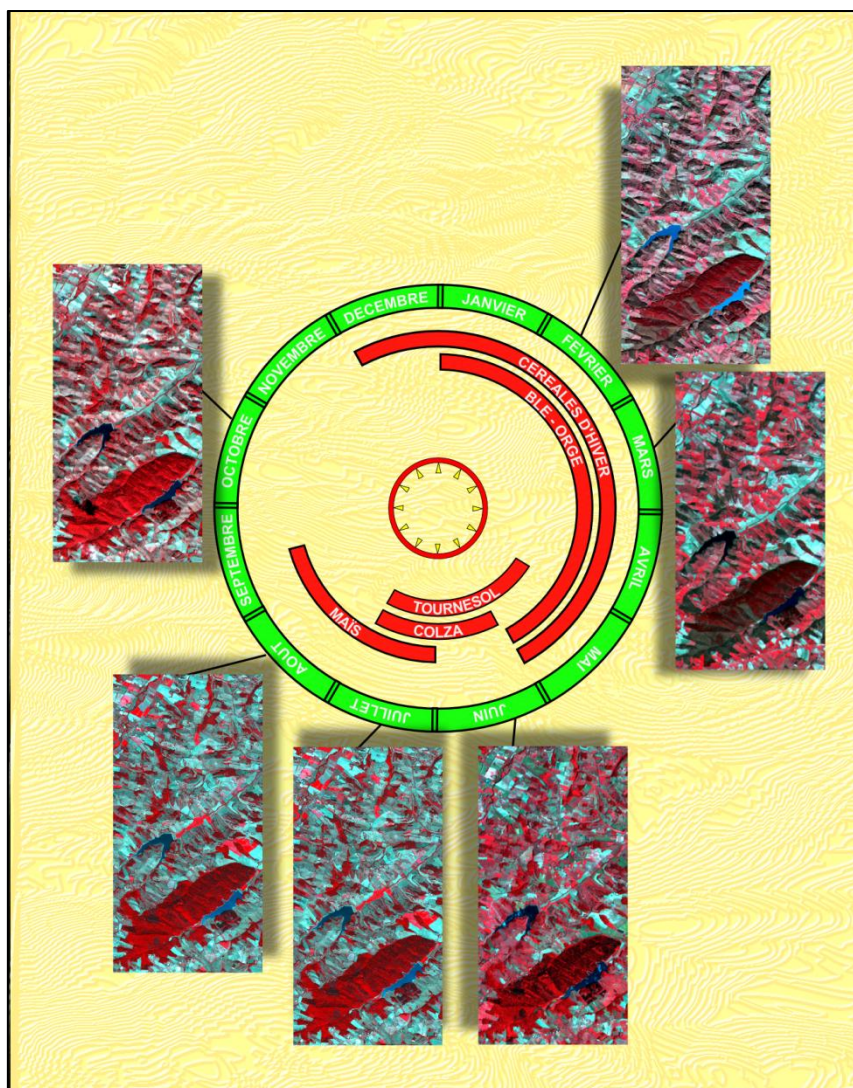


Figure 5 : Représentation du calendrier culturel des principales cultures présentes dans les Coteaux de Gascogne associé à différentes prises de vue SPOT en 2009.

- Sur la base de ce calendrier culturel et des cultures retrouvées dans les coteaux de Gascogne, un métafichier regroupant des images SPOT 4 et SPOT 5 de l'année 2009⁴ a été réalisé par Marcela Arias et Nicolas Barry dans le cadre d'un mini-projet SIGMA suite à un rééchantillonnage des différentes images à dix mètres de résolution (Arias et Barry, 2011). Ce métafichier contient les informations contenues dans les canaux de sept images dont le récapitulatif apparaît dans le tableau 1.

⁴ 2009 est l'année la plus récente pour laquelle DYNAFOR dispose d'un lot d'images importants pour sa zone des Coteaux de Gascogne. Ces images sont issues du programme Kalidéos OSR MiPy: ce programme du CNES met gratuitement à disposition de la communauté scientifique des bases de données de télédétection de référence. Dynafor en tant que membre fondateur de l'Observatoire Spatial Régional peut en bénéficier pour son site d'études LTER Vallées et Coteaux de Gascogne.

Date de prise de vue	Satellite	Résolution (m)	Nombre de bandes spectrales
04/02/2009	SPOT 5	10	4
24/06/2009	SPOT 4	10	4
31/07/2009	SPOT 5	10	4
15/08/2009	SPOT 4	10	4
19/08/2009	SPOT 5	10	4
31/08/2009	SPOT 5	2,5	3
11/09/2009	SPOT 4	20	4

Tableau 1: Récapitulatif des caractéristiques des images SPOT retenues pour le métafichier.

L'emprise de ce métafichier recouvre une zone de 1820,6 km² (figure 6).

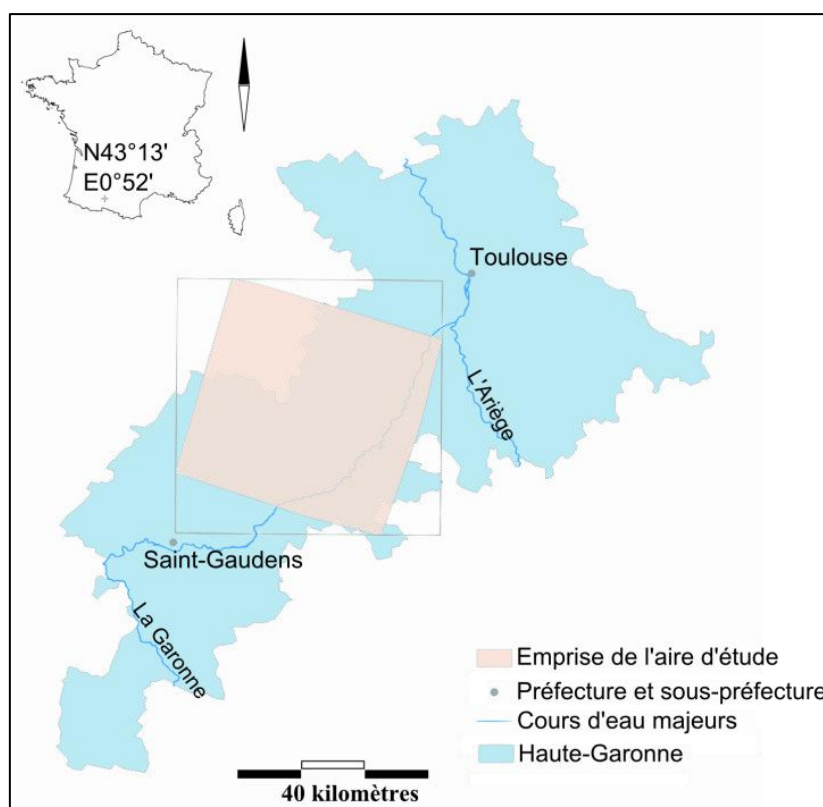


Figure 6 : Représentation de l'aire d'étude et de la surface recouverte par le métafichier.

- D'autres données géographiques supplémentaires, dites de deuxième ordre ont été mobilisées, à partir du Système d'Information existant⁵, au cours du stage pour compléter. Il s'agit des photos aériennes de 2006 de la BD ORTHO de l'IGN, les données vectorielles (limite administrative, réseaux routiers et hydrographiques) de la BD TOPO des départements étudiés.

3) Présentation générale de la chaîne de traitement.

Tout notre travail s'appuie sur des recherches effectuées au Canada sous la direction de Lenore Fahrig, et plus particulièrement sur un article en cours de parution qui décrit la

⁵ DynaGIS : base de données géographiques construite sur le terrain des Coteaux de Gascogne regroupant des données sur le paysage pour l'étude des relations paysage et activités agricoles et forestières et biodiversité)

méthode cartographique en détaillant ses aspects techniques et en justifiant les paramètres choisis (Pasher *et al.*, draft). Le but de mon travail de stage va donc être d'établir des cartes d'hétérogénéité du paysage et de réaliser un plan d'échantillonnage préalable aux campagnes de terrain (récolte des données biologiques), en tenant compte des spécificités du terrain d'étude (Europe versus Canada) et de la formalisation des besoins des thématiciens écologues ou agronomes, experts de la zone d'étude. Pour cela, après avoir choisi les données de départ, nous avons défini différentes étapes nécessaires pour réappliquer la méthode développée par Pasher *et al.* à d'autres zones:

- 1) Classification de l'occupation du sol en concertation avec les chercheurs pour fixer les types à conserver et pouvoir renseigner la composition du paysage ;
- 2) Segmentation vectorielle des limites de parcelles pour pouvoir renseigner la configuration du paysage ;
- 3) Prétraitement à la réalisation de cartes d'hétérogénéité du paysage par analyse par fenêtre mobile ;
- 4) Réalisation des cartes d'hétérogénéité du paysage par des analyses par fenêtre mobile ;
- 5) Echantillonnage des points de suivis faunistiques et floristiques.

La figure 7 montre ces différentes étapes ainsi que les différents logiciels intervenant dans chacune de ces étapes.

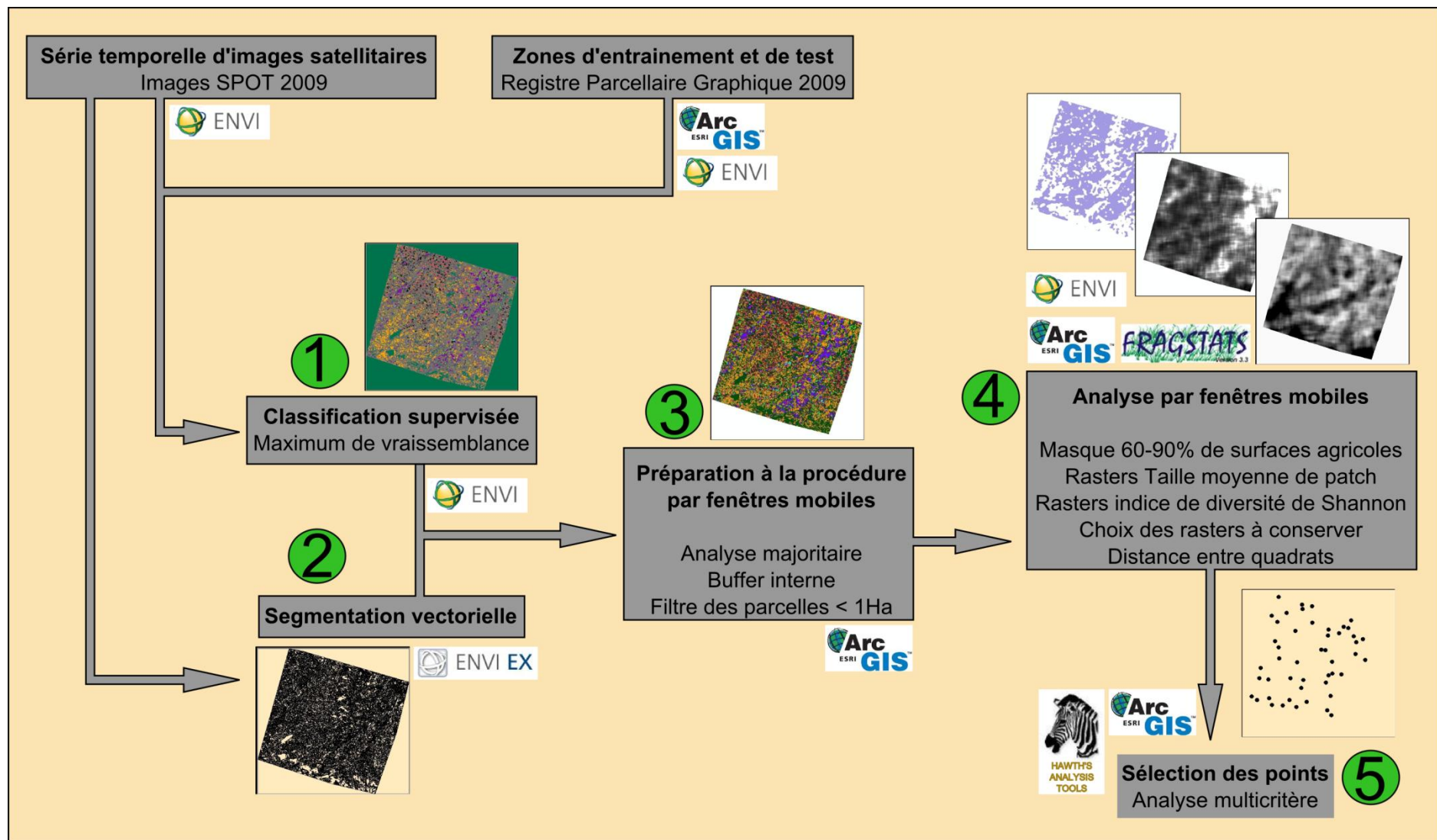


Figure 7 : Vue générale et chronologique de la chaîne de traitement et logiciels utilisés pour chaque étape.

4) Plateforme logicielle.

Le parcours de la chaîne de traitement implique l'utilisation de différents logiciels entre chaque étape mais aussi au sein de certaines d'entre elles (Figure 7).

Le logiciel de télédétection ENVI v4.7 (ITT, 2009) a été utilisé pour réaliser les classifications supervisées ainsi que pour calculer des statistiques spatiales suite au calcul des indices d'hétérogénéité du paysage. Par ailleurs, l'extension ENVI EX ainsi que le logiciel Definiens Professional v5.0 ont été utilisés pour effectuer les procédures de segmentation vectorielle (orientée objet).

ArcGIS v9.3.1 (ESRI, 2006) a été utilisé pour préparer la constitution des régions d'entraînement et de test relatives aux classifications, pour la réalisation de digitalisation manuelle de certaines parcelles, pour la préparation aux analyses par fenêtres mobiles, pour la réalisation des masques comprenant entre 60 et 90% de surface agricole ainsi que pour la mise en œuvre de ces masques sur les rasters basés sur les indices d'hétérogénéité. Ce logiciel a également été mis à contribution lors de la phase de sélection des points d'échantillonnage avec son module *Hawths Tools* v3.27, (Beyer 2004).

FRAGSTATS v3.3 (McGarigal *et al.*, 2002) a servi pour le calcul de rasters basés sur différentes métriques paysagères utilisant une procédure de fenêtre mobile. Il aura notamment permis le calcul des rasters basés sur les indices d'hétérogénéité du paysage ainsi que le calcul des rasters basés sur le pourcentage de surface agricole.

Les différents graphiques bivariés ont été réalisés avec le tableur Excel v2007 ainsi qu'avec le logiciel open source R v2.13.0. Ce dernier logiciel a également été utilisé pour le calcul des différentes analyses statistiques.

V. Etape 1 : Classification de l'occupation du sol pour définir la composition du paysage étudié.

Après la phase de sélection des données de base, il faut déterminer les différentes modalités d'occupation du sol. C'est en effet sur ce socle que les analyses de description du paysage qui nous serviront d'indices d'hétérogénéité pourront être effectuées. Afin de déterminer cette occupation du sol, une classification supervisée a été réalisée avec le logiciel ENVI 4.7 en utilisant des zones d'apprentissage et de test issu de données agricoles (Registre Parcellaire Graphique) et de données vectorielles digitalisées sur photo aérienne pour des classes non agricoles.

1) Utilisation de données agricoles du Registre Parcellaire Graphique (RPG) pour réaliser des échantillons d'apprentissage et de test de la classification.

Afin d'obtenir des échantillons d'apprentissage et de test pour la réalisation et la validation de la classification supervisée, nous avons choisi d'utiliser des îlots homogènes du RPG. Cette approche nous a ainsi permis d'obtenir facilement un grand nombre de régions d'intérêt (ROI) pour la phase de classification. Initialement, différentes modalités d'occupation du sol sont inscrites dans ce registre comme cela apparait dans le tableau 2, mais

toutes ne vont pas correspondre à des cultures retrouvées dans les Coteaux de Gascogne. D'autres pourraient être présentes dans la zone d'étude mais en quantité anecdotique et il nous a semblé nécessaire d'effectuer un recodage de ces modalités. Ce travail a été réalisé dans un premier temps dans le cadre d'un mini-projet SIGMA (Arias et Barry, 2011). Dans un second temps, nous avons repris ce recodage afin de réduire le nombre de classes agricoles pour ne garder que les cultures majoritaires afin d'améliorer les performances de la classification. En outre, il a fallu compléter le jeu de données ROI pour les types non-agricoles que nous souhaitons retrouver dans la classification à savoir les bois, les haies, le bâti. Ces différents recodages successifs apparaissent dans le tableau 2 ci-dessous :

Classes du RPG	Classes retenues dans un premier temps par Arias et Barry	Classes retenues dans un second temps par Arias et Barry	Classes retenues dans notre travail
Blé tendre	Blé tendre	Blé/Orge/Autres céréales	Céréales
Orge	Orge		
Autres céréales	Autres céréales		
Colza	Colza	Colza	Colza
Maïs	Maïs	Maïs	Maïs
Surface sans production	Surface sans production	Prairies/Légumineuses à grain/Surfaces sans production	Prairies
Prairies permanente	Prairies permanente		
Prairies temporaires	Prairies temporaires / légumineuse à grain		
Légumineuses à grain			
Autres Oléagineux	Autres Oléagineux	Autres Oléagineux	Protéo-Oléagineux
Protéagineux	Protéagineux	Protéagineux	
Tournesol	Tournesol	Tournesol	Tournesol
Plantes à fibres	Plantes à fibres	Plantes à fibres	x
semences	x	x	x
Gel industriel	x	x	x
Autres gels	x	x	x
Riz	x	x	x
Fourrage	x	x	x
Estives Landes	x	x	x
Vergers	Vergers/arboriculture	Vergers/arboriculture	x
Arboriculture			x
Vignes	Vignes	Vignes	x
Fruits à coque	x	x	x
Oliviers	x	x	x
Autres cultures industrielles	x	x	x
Légumes-fleurs	x	x	x
Canne à sucre	x	x	x
Divers	x	x	x

Tableau 2 : Modalités d'occupation du sol inscrites dans le Registre Parcellaire Graphique et les différents recodages successif réalisés. Les croix illustrent les classes non prises en compte du fait de leur inexistence dans l'aire d'étude ou de leur rareté.

2) Une étude préliminaire.

Cette étude préliminaire a été réalisée dans le cadre d'un mini-projet SIGMA (Arias et Barry, 2011). Ce travail a permis de préparer notre étude en mettant en forme les données du RPG afin de ne retenir que les îlots homogènes. Un échantillonnage sur ces îlots a permis de créer deux fichiers *vecteur* qui ont servi de base pour l'un d'apprentissage et pour l'autre de validation de la classification.

Une première classification selon un algorithme du maximum de vraisemblance a ainsi été calculée avec une précision globale de 57,27%. Sur la base de la matrice de confusion de cette classification et grâce à l'indice de Jeffries-Matusita permettant d'identifier la séparabilité de deux modalités d'occupation du sol, Arias et Barry en sont venus à proposer le regroupement de différents types de cultures afin d'améliorer la précision de la classification. De cette manière et après discussion avec les chercheurs agronomes, le blé tendre, l'orge et les autres céréales ont été regroupés sous le label « céréales » ; les prairies permanentes et temporaires ont été recodées en sous le terme générique de « prairies » auxquelles ont été ajoutées les surfaces dites gelées et les légumineuses ; les deux types de friches (arborée et à genévrier) ont également été regroupées sous le label « Friche ».

Ce recodage a alors permis à Arias et Barry d'améliorer de manière notable la qualité de leur classification en atteignant une précision globale de 72,91%. Une double analyse majoritaire a alors permis d'éliminer les pixels isolés et de gagner quelques pourcentages de précision globale en atteignant une valeur de 76,03% soit 18,76% de mieux que la première classification pour une classification comportant 15 classes d'occupation du sol.

3) Discussion des nouveaux résultats de la classification.

Suite à ce travail d'Arias et Barry, nous avons choisi de reprendre le travail de classification pour réduire le nombre de classes en se focalisant sur les classes agricoles majoritaire et en regroupant dans une classe unique « contexte non agricole » les classes non agricoles afin de respecter le protocole de Pasher *et al.* et d'améliorer *in fine* la précision globale. Nous avons alors effectué un nouveau recodage des différentes modalités d'occupation du sol afin de n'en garder que sept pour lesquelles nous disposerons de régions d'intérêt. Comme le rappelle le tableau 2, les « protéagineux » et les « autres oléagineux » ont été rassemblés ; les « plantes à fibres », les « vignes » et les « vergers et arboricultures » étant particulièrement rares sur notre aire d'étude, ces modalités ont été éliminées et associées à une modalité « Contexte non-agricole ». Cette dernière modalité contenait également le bâti, les surfaces en eau ainsi que les forêts et les friches.

Suite à ce recodage, nous avons effectué la classification selon un algorithme de maximum de vraisemblance comme précédemment. La séparabilité des régions d'apprentissage a été testée par l'indice de Jeffries-Matusita. Cet indice varie de 0 à 2 et la séparabilité des classes est considérée comme bonne à partir d'une valeur de 1,9. Les résultats sont rassemblés dans le tableau 3.

	Colza	Maïs	Prairies	Proteo-Oléagineux	Tournesol	Contexte non-agricole
Maïs	1,998					
Prairies	1,931	1,991				
Proteo-Oléagineux	1,920	1,958	1,941			
Tournesol	1,963	1,986	1,941	1,919		
Contexte non agricole	1,995	1,993	1,941	1,993	1,985	
Céréales	1,744	1,951	1,869	1,721	1,828	1,990

Tableau 3 : Matrice des indices de Jeffries-Matusita relatifs à la séparabilité des différentes modalités d'occupation du sol suite à la classification par le maximum de vraisemblance.

Dans l'ensemble, la séparabilité des zones d'apprentissage est plutôt bonne à quelques exceptions près. Compte tenu du fait que certaines de ces classes sont un assemblage hétéroclite de différentes cultures (les classes Protéo-Oléagineux et Céréales), ces valeurs localement faibles pourraient trouver une explication.

Cette classification a été validée par une matrice de confusion réalisée sur la base des échantillons tests pour chaque modalité d'occupation du sol. Nous avons obtenu un taux de précision global de 81%. La matrice de confusion correspondante apparaît dans le tableau 4 suivant :

Classes Test	Modalités d'occupation du sol de la classification						
	Cereales	Colza	Maïs	Prairies	Proteo-oleagineux	Tournesol	Contexte non-agricole
Cereales	84.64	17.95	4.08	1.64	9.31	4.61	1.01
Colza	7.71	62.93	0.08	3.91	1.22	1.8	0.06
Maïs	0.05	0.26	84.87	0.55	9.11	0.7	0.74
Prairies	5.66	13.27	3.17	90.7	6.83	4.26	3.85
Proteo-oleagineux	0.92	1.79	4.62	0.43	68.58	2.86	0.31
Tournesol	0.95	3.18	2.69	1.1	4.38	85.07	16.23
Contexte non agricole	0.06	0.61	0.47	1.67	0.56	0.69	77.8

Tableau 4: Matrice de confusion obtenue suite à la classification par le maximum de vraisemblance. Dans la diagonale (en jaune figurent les pourcentages de bien classés et les erreurs supérieures à 5% ont été surlignées en rose).

Bien que relativement faible (il y a quand même près de 20% d'erreur), le niveau de précision globale est équivalent à celui trouvé par Pasher *et al.* pour un paysage nettement plus simple à classer (5 classes retenues : Prairie ; Soja ; Maïs, Céréales ; Non-agricole). En effet, comme le montre la figure 8 issue de cet article, les grandes parcelles cultivées du continent nord-américain présentent des formes plus géométriques que les paysages morcelés des coteaux de Gascogne et une diversité de cultures plus faible (4 modalités de cultures contre 7 dans notre cas). Par ailleurs, dans le cadre d'une étude à petite échelle comparable, Peña-barragan *et al.* (2011) présentent une méthode de segmentation pour identifier des modalités d'occupation du sol et trouve une précision globale (*Global accuracy*) proche de celle que nous obtenons. La précision utilisateur par modalité (*User accuracy*) est parfois

même moins bonne que les plus mauvaises des nôtres. D'après Lenore Fahrig (communication personnelle), il est évident que le niveau de précision global maximal est souhaitable mais que de toute façon, les phases de terrain ultérieures permettront de corriger les limites atteintes dans la phase de classification et de segmentation. Sur cette base, nous avons considéré que cette étape du protocole était accomplie.

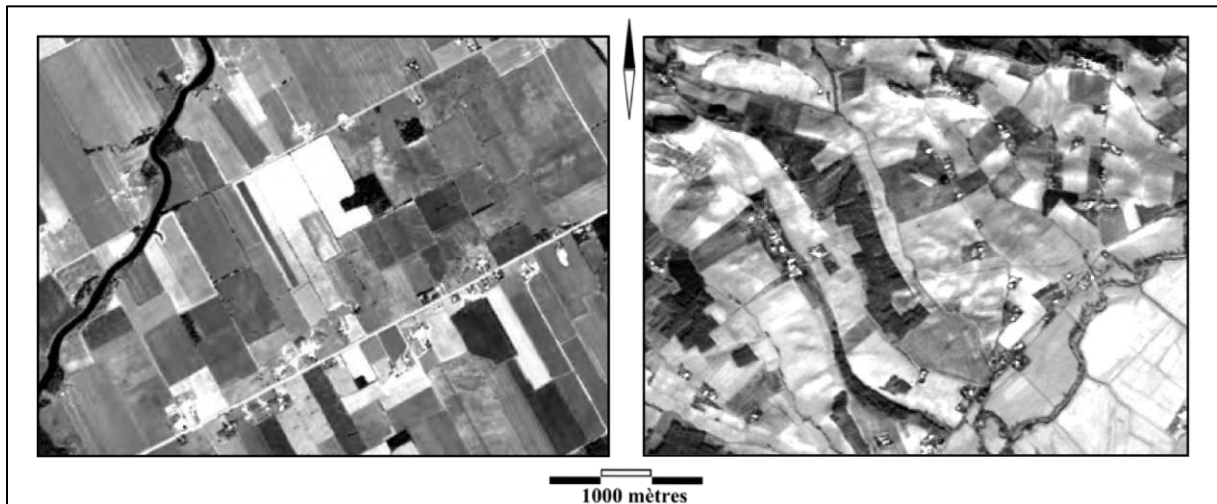


Figure 8 : Image panchromatique SPOT issue du travail de Pasher *et al.* . On y distingue aisément la structure très géométrique du paysage agricole qui contraste fortement avec les paysages très découpés rencontrés dans les coteaux de Gascogne vus sur les figures précédentes.

Il est important de rappeler que le but de ce travail est avant tout exploratoire et vise à adapter une chaîne de traitement mise au point pour décrire un paysage agricole nord-américain à différents paysages européens possédant chacun ses spécificités. Ainsi, il est prévu de recommencer toute la chaîne de traitement durant l'automne 2011 avec le lancement officiel du projet FARMLAND sur tous les terrains européens avec des images satellites actualisées. Nous pouvons d'ores et déjà proposer différentes pistes afin d'améliorer la classification de l'occupation du sol :

- Il est envisageable d'effectuer des masques successifs de différentes parties de l'image satellite comme les parties boisées, les cours d'eau et le bâti. En effet, une part importante des erreurs de classement du contexte non agricole est liée à une confusion de cette modalité avec le tournesol (probablement due aux friches). L'utilisation de prises de vues tardives durant l'automne et /ou l'hiver pourrait ainsi permettre de lever cette ambiguïté en identifiant mieux le contexte non agricole et en réalisant un masque.
- Les parcelles du RPG présentent des limites manque parfois de précision et peuvent recouvrir les occupations du sol voisines (forêts, haies...). L'application à ces parcelles d'une zone tampon interne dans la réalisation des échantillons d'apprentissage et de validation pourrait ainsi permettre de gagner en capacité de discrimination.

VI. Etape 2 : Segmentation vectorielle des limites de parcelles pour définir la configuration du paysage étudié.

1) Introduction aux méthodes de segmentation.

L'étape précédente nous a permis d'obtenir une classification de l'occupation du sol selon les différentes modalités que nous avons retenues. Toutefois, cette occupation du sol n'intègre pas les limites entre les parcelles. Ainsi, pour plusieurs parcelles contiguës ayant un même type de culture, la classification n'identifiera qu'une seule et même parcelle. Cela risquerait de nous poser problème pour la suite du protocole dans la mesure où cela interférerait directement sur les calculs des deux indices d'hétérogénéité que nous avons choisis. Pour remédier à cela et obtenir les limites de chaque parcelle d'utilisation agricole, nous avons utilisé une procédure de segmentation vectorielle. Cette approche orientée-objet va utiliser différentes caractéristiques spatiales, spectrales et de texture d'une image multispectrale ou panchromatique.

Pasher *et al.* ont utilisé le logiciel Definiens Professionnal v5.0 pour effectuer la segmentation. Nous disposons d'une licence pour ce logiciel à l'ENSAT, mais un module d'ENVI appelé ENVI EX permet également d'effectuer cette opération avec une interface plus ergonomique. Nous avons donc comparé l'usage de ces 2 outils.

Une segmentation sous ENVI EX se fait en deux phases. La première phase découpe le paysage en un grand nombre de parcelles et la deuxième agrège ces parcelles en entités de plus grande taille. Pour chaque phase on doit choisir un paramétrage allant de 0 à 100 (paramétrage d'échelle pour la première phase et paramétrage de fusion pour la deuxième). Avec un paramétrage d'échelle faible il y aura beaucoup de petites parcelles ; avec un paramétrage de 100 il n'y en aura plus qu'une. Le paramétrage de fusion permet de compléter et d'affiner l'étape précédente.

Avec le logiciel Definiens Professionnal v5.0, différents algorithmes de segmentation sont proposés. Nous avons choisi d'utiliser celui de la segmentation multi-résolution en fixant le paramètre d'échelle à 60, valeur qui avait donné une classification intéressante avec ENVI EX et qui était utilisée par Pasher *et al.* .

2) Résultats et validation des segmentations.

Afin de valider la segmentation, Pasher *et al.* proposent d'effectuer deux tests :

- Sélection d'une centaine de parcelles délimitées automatiquement (par la segmentation) et manuellement (par digitalisation) et étude de la corrélation entre les surfaces de parcelles obtenues par les deux approches.
- Sélection aléatoire d'une quinzaine de carrés de 1 km de côté dans l'image obtenue et étude de la corrélation entre le nombre de parcelles obtenues par les deux

approches (automatique et manuelle) ainsi que pour leur surface. Ces quadrats d'échantillonnage apparaissent sur la figure 9:

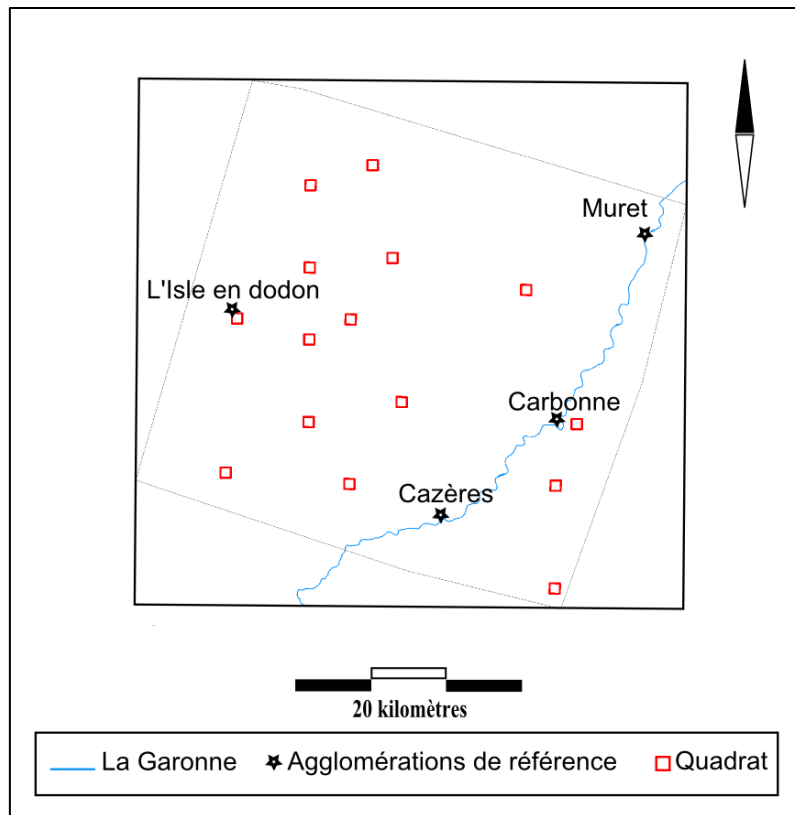


Figure 9 : Distribution des quadrats d'échantillonnage en vue de tester la corrélation entre les digitalisations manuelle et automatique.

Etant donné qu'ENVI offre une grande souplesse dans les possibilités de paramétrage, nous avons effectué une série de segmentations dont nous avons évalué la validité selon la deuxième approche proposée par Lenore Fahrig (les 15 carrés de 1 km²). Cette démarche nous a ainsi permis de retenir la meilleure des segmentations réalisées.

Dans un premier temps et afin de gagner du temps, nous avons effectué ces segmentations à partir d'une image simplifiée constituée de 2 bandes, XS1 (le vert) et XS 2 (le rouge) présentant un meilleur contraste que les bandes infrarouges pour la date du 24 juin 2009. Cette date a été choisie comme intermédiaire entre l'été et l'hiver et comme pouvant regrouper des stades relativement avancés des différentes cultures présentes sur la zone.

Dans un deuxième temps, nous avons effectué cette analyse à partir de l'image contenant toutes les bandes disponibles pour toutes les dates du métafichier. Les calculs sont alors beaucoup plus lourds (à peu près une heure de calcul par segmentation) mais la précision a été considérablement augmentée (coefficient de corrélation de 0.72 contre 0.54 précédemment). La figure 10 montre la distribution de ces différents coefficients de corrélation.

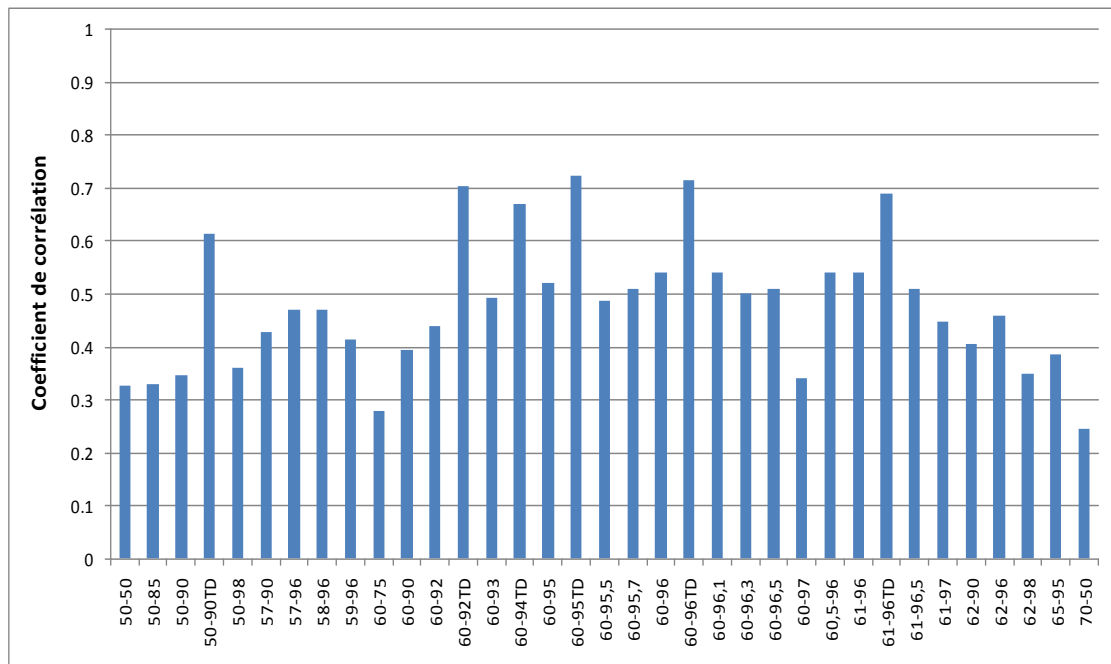


Figure 10 : Distribution des coefficients de corrélation de Pearson entre le nombre de parcelles obtenues par segmentation automatique et par digitalisation manuelle.. Le suffixe TD illustre les calculs réalisés sur une image de départ rassemblant toutes les bandes spectrales disponibles et toutes les dates. Les deux nombres correspondent au paramètre d'échelle pour le premier et au paramètre de fusion pour le second.

La figure 10 montre dans quelle mesure le fait d'utiliser toutes les bandes disponibles pour toutes les dates améliore la qualité de la segmentation. Par ailleurs, la segmentation réalisée avec Definiens Professionnal 5.0 a également donné de bons résultats sur la base de l'analyse de corrélation entre le nombre de parcelles retrouvées automatiquement et celui obtenu par digitalisation manuelle (corrélation de 0,7). Toutefois la forme des parcelles semble moins cohérente que pour celles obtenues par ENVI. La comparaison entre les deux logiciels apparaît sur la figure 11. Etant donné que la corrélation est très légèrement supérieure avec ENVI, nous avons conservé la meilleure segmentation obtenue avec ce logiciel soit celle correspondant aux paramètres d'échelle à 60 et de fusion à 92 pour toutes les dates et toutes les bandes.

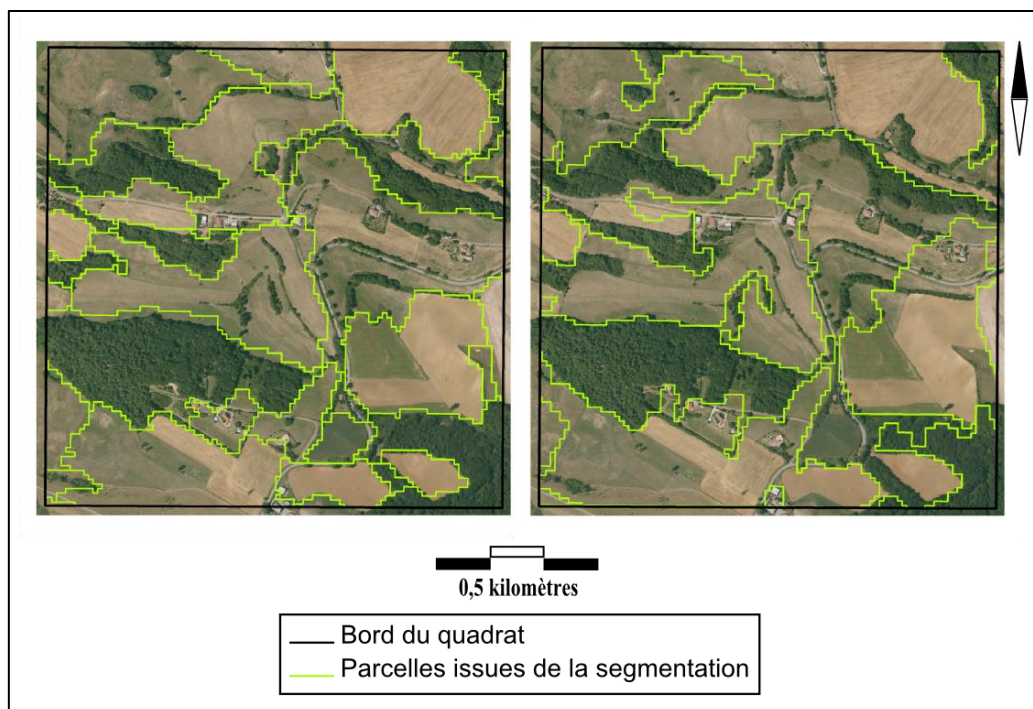


Figure 11 : Superposition de photographies aériennes et de représentation cartographique de la segmentation et.
 a) A gauche : segmentation avec ENVI EX pour un paramétrage 60-92 pour toutes les bandes disponibles et toutes les dates. b) A droite : segmentation avec Definiens Professionnal 5.0 pour un paramétrage 60 pour toutes les bandes disponibles et toutes les dates.

3) Discussion des résultats de la segmentation.

Les différentes analyses de segmentation vectorielle que nous avons réalisées nous ont permis d'atteindre un résultat acceptable selon les critères de Lenore Fahrig. Pasher *et al.* (*op. cit.*), atteignent une corrélation de 0.83 pour le nombre d'entités retrouvées dans les 15 quadrats de manière automatique et par digitalisation manuelle. Il peut être opportun de rappeler la différence qui existe entre les paysages agricoles nord-américains et ceux que l'on retrouve dans les Coteaux de Gascogne ayant servi de base à notre étude. La figure 8 tirée du travail de Pasher *et al.* permet de mieux comprendre le décalage qui existe entre ces systèmes de production. Cette complexité de nos paysages est probablement l'un des éléments ayant contribué à ce que nos corrélations entre la segmentation et la digitalisation manuelle soit plus faibles que ce qui avait été observé au Canada. Le fait que ces auteurs obtiennent ce résultat avec une seule image panchromatique alors que l'ajout des canaux et d'images de dates différentes dans notre cas ait amélioré nos statistiques va également dans ce sens.

L'utilisation dans notre cas de la surface moyenne des parcelles plutôt que leur nombre ne nous a pas semblé être un élément de validation intéressant. En effet avec cette variable, la corrélation a toujours été très forte (supérieure à 0,9) et les changements de paramétrage des segmentations n'avait d'effet que sur la troisième décimale...

Encore une fois, nous rappellerons que le but de ce travail est avant tout exploratoire et visait à adapter une méthodologie mise au point pour des paysages bien différents des nôtres. Cette étape de la chaîne de traitement sera, elle aussi, à remettre en œuvre lorsque le projet

FARMLAND sera lancé à l'automne 2011. Notamment, cette phase de segmentation devra, elle aussi, être effectuée de nouveau sur des images satellites plus récentes.

VII. Etape 3 : Prétraitement pour l'analyse par fenêtre mobile avec FRAGSTATS v3.3.

1) Présentation du logiciel FRAGSTATS v3.3.

Ce logiciel couramment utilisé en écologie du paysage nous a permis de calculer de nouveaux rasters sur la base de la carte d'occupation du sol réalisée lors des deux étapes précédentes par une procédure dite « par fenêtre mobile ». Concrètement, la procédure de fenêtre mobile se comporte un peu comme un filtre de lissage. Une fenêtre (de forme ronde ou carré dont la taille⁶ est à déterminer) parcourt le paysage et va donner une valeur au pixel central de la fenêtre qui dépendra de la métrique choisie (valeur moyenne, médiane, indice de diversité...). Parmi les indices et les métriques que ce logiciel va permettre de calculer, l'indice de diversité de Shannon est un indice de l'hétérogénéité de composition alors que la taille moyenne des taches (appelés aussi *patches*) de végétation est un indice de l'hétérogénéité de configuration (Fahrig *et al.*, 2010). Ces deux indices sont présentés dans les annexes. L'utilisation de la procédure par fenêtre mobile de FRAGSTATS v3.3, innovante pour l'UMR DYNAFOR a donné lieu à la réalisation d'une fiche qualité stockée dans le gestionnaire de documents qualité de l'unité et décrivant sa mise en œuvre technique. Cette fiche figure dans les annexes de ce mémoire.

2) Préparation aux traitements par fenêtre mobile.

Lors de chacune des étapes précédentes (classification et segmentation), nous avons vu que des phénomènes assimilés à des erreurs pouvaient être observés. Ces erreurs peuvent, dans le cas de la classification, être reliées à différentes causes comme par exemple la présence de sol nu au milieu d'une parcelle en végétation. Aussi au cours de la segmentation, cela pourrait aboutir à la présence d'entités de petite taille n'ayant pas de pertinence à notre niveau d'analyse et venant biaiser nos calculs ultérieurs. Avant d'utiliser FRAGSTATS, différents traitements préliminaires successifs ont donc été appliqués à nos données d'occupation du sol :

- Analyse majoritaire de la segmentation renseignée par la classification.
- Elimination des parcelles dont la surface était inférieure à 1 hectare (recodées en contexte non agricole).
- Mise en place d'une zone tampon interne de 8 mètres (recodée elle aussi en contexte non agricole).

⁶ Ici la taille est appelée « radius » dans Fragstats v3.3 et correspond à la plus courte distance entre la cellule centrale de la fenêtre et le bord de celle-ci. Ainsi, la longueur d'une fenêtre carrée sera égale à la somme de deux fois le radius et de la taille du pixel central.

Les deux premiers traitements nous ont ainsi permis de limiter considérablement les artefacts de la classification et de la segmentation. Le *buffer* interne de 8 mètres visait plus spécifiquement à limiter les effets d'interaction entre les parcelles contiguës (effets des apports d'intrants dans les parcelles cultivées notamment). La figure 12 montre le résultat de ces trois phases.

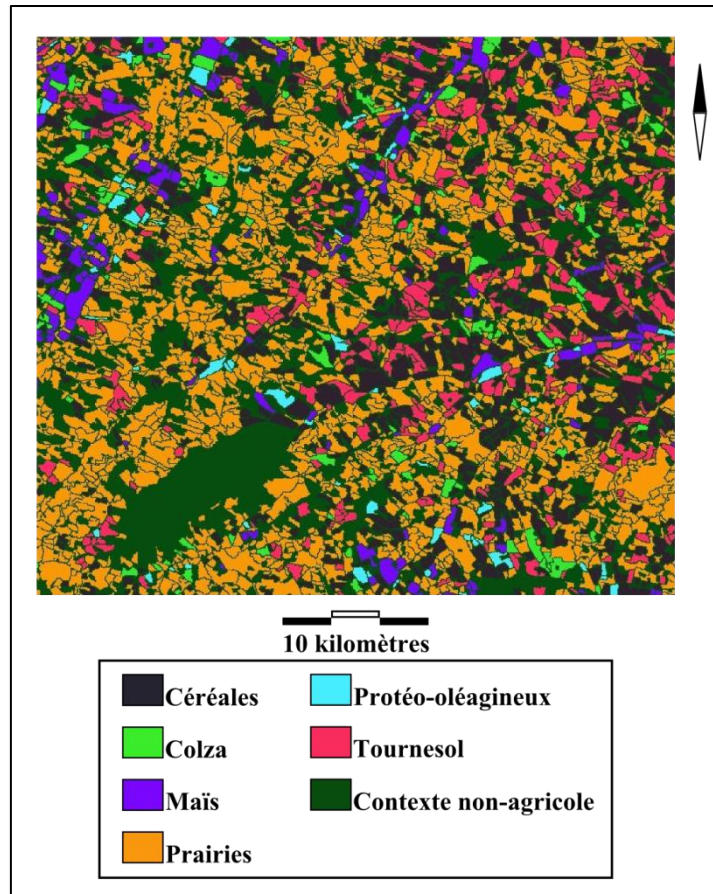


Figure 12 : Détail de la carte d'occupation du sol selon les sept modalités retenues après analyse majoritaire de la segmentation renseignée par la classification, recodage des parcelles ayant une surface inférieure à un hectare en contexte non-agricole et mise en place d'une zone tampon interne de huit mètres (également recodée en contexte non-agricole).

Il est possible de recoder le compte numérique de ces pixels correspondant à du contexte non-agricole afin qu'ils ne participent pas aux calculs ultérieurs. Ces pixels recevront une valeur particulière (classiquement 99) comme compte numérique pour être identifiés. Il sera alors possible soit d'intégrer ces pixels dans les calculs ultérieurs comme données manquantes (compte numérique positif +99) soit de les ignorer (compte numérique négatif -99). Ce procédé permettra également d'éliminer d'éventuels effets de bord en créant une bordure à l'image.

La dernière étape de ces traitements préliminaires a consisté à convertir le raster d'occupation du sol au format ASCII, ce dernier étant un des formats d'entrée dans le logiciel FRAGSTATS v3.3. Pour se faire nous avons utilisé les scripts internes à Arcgis : « Entités vers raster » puis « raster vers ASCII » de la *toolbox* : « Outils de conversion ». A ce stade,

une intervention manuelle est à faire sur le fichier ASCII à qui on doit retirer l'en-tête sous un éditeur de texte (4 premières lignes)⁷. Malgré l'amélioration de l'interopérabilité entre les logiciels, des problèmes de compatibilité de fichiers existent toujours et sont à résoudre manuellement. Cet en-tête est à conserver en vue de l'import, dans ArcGis des nouveaux rasters basés sur les indices d'hétérogénéité (« ASCII to Raster » située dans le menu « Outils de conversion »).

VIII. Etape 4 : Analyses par fenêtre mobile.

Afin de dissocier les effets éventuels des politiques agricoles sur la biodiversité de ceux des autres composantes de l'espace rural, nous avons retiré le contexte non-agricole des calculs de rasters basés sur des indices d'hétérogénéité en leur affectant un code de données manquantes. Afin de détecter ces effets, Fahrig recommande de sélectionner les points d'échantillonnage sur lesquels effectuer les relevés de biodiversité dans les zones comportant entre 60 et 90% de surface agricole. Une de nos premières tâches a donc été de calculer un masque retenant cet intervalle de valeurs.

Puis nous avons calculé les rasters basés sur les indices d'hétérogénéité de composition et de configuration, respectivement l'indice de diversité de Shannon et la taille moyenne des taches de végétation. Comme cela a été exposé dans le paragraphe précédent, le calcul de ces rasters se fait à l'intérieur d'une fenêtre mobile parcourant l'image d'occupation du sol dont la taille est paramétrable. Ainsi, différentes échelles d'analyses peuvent être envisagées. Selon Pasher *et al.*, les effets de la structure du paysage apparaissent souvent plus fortement à des échelles différentes voisines de celle à laquelle les suivis de biodiversité sont menés, raison pour laquelle différentes échelles d'analyse sont fréquemment envisagées. Après avoir calculé un ensemble de rasters construits sur la base de fenêtres mobiles de différentes tailles, il a donc fallu déterminer à quelles échelles il pourrait être intéressant de travailler.

A terme, les quadrats d'échantillonnage devant être distribués géographiquement dans l'image de façon à être le moins corrélés spatialement entre eux (voire pas du tout dans l'idéal), il a fallu déterminer la distance entre le barycentre de deux quadrats à partir de laquelle cette autocorrélation était la plus faible voire pouvait être considérée comme nulle.

1) Détermination des différentes échelles d'analyse.

Pasher *et al.* ont choisi des fenêtres mobiles dont la taille pouvait aller de 500 mètres à 7,5 kilomètres (0,5, 1, 2, 3, 4, 5, 6 et 7,5 km). Compte tenu du temps considérable que FRAGSTATS v3.3 peut prendre pour calculer un raster sur une image de départ de 1820 km² (de quelques heures pour l'étendue la plus petite à quelques jours pour la plus grande), et compte tenu de la complexité de configuration de nos paysages, des étendues de 0,25 à 3 km ont été préférées dans notre cas d'étude en augmentant progressivement la taille de ces fenêtres (0,25, 0,35, 0,5, 0,75, 1, 1,5, 2 et 3 km).

⁷ Extrait de l'aide en ligne de Fragstats sur ce point : ASCII file, no header. Each record should contain 1 image row. Cell values should be separated by a comma or a space(s). Note, it will be necessary to strip (delete) the header information from the image file if it exists, but be sure to keep it for later reference regarding background cell value, # rows, and # columns

2) Calcul des masques retenant entre 60 et 90 % de surface agricole.

Le protocole requiert tout d'abord le calcul d'un masque retenant les zones comprenant de 60 à 90% de surface agricole pour chaque étendue. Cette procédure se fait sous FRAGSTATS v3.3 avec des fenêtres mobiles avec la fonction PLAND qui permet de calculer le pourcentage de surface du paysage pour une classe donnée. Pour cela nous avons reclassé le raster de départ de manière binaire (agricole et non agricole). La figure 13 montre les masques obtenus pour les étendues de 0,25, 1 et 3 kilomètres. On notera qu'à partir de 3 km d'étendue, la surface de zone agricole comprise entre 60 et 90% baisse sensiblement et qu'un effet de bordure est aisément identifiable. En effet, lorsque FRAGSTATS v3.3 effectue ses calculs par fenêtre mobile, il ne peut travailler que sur des pixels pouvant être au centre d'une fenêtre (voir la fiche qualité relative à l'utilisation de ce logiciel en annexe). Autrement dit, plus la fenêtre est grande plus celle-ci se décale des bords de l'image initiale (effet de rognage de la bordure) et plus le raster en sortie est de taille réduite par rapport à l'image de départ comme cela se remarque sur la figure 13.

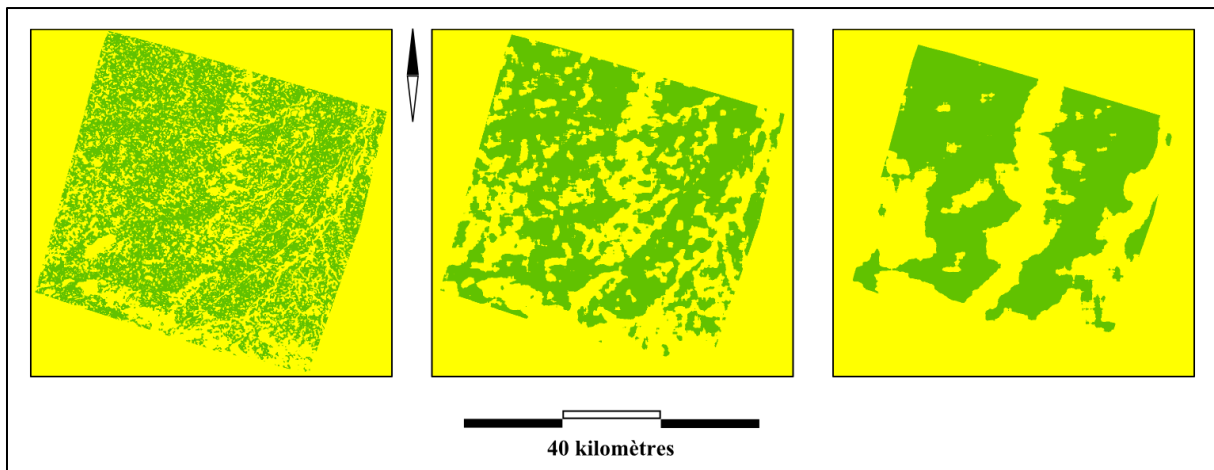


Figure 13 : Rasters retenant les pixels centraux des fenêtres mobiles de 0,25, 1 et 3 kilomètres pour lesquelles la fenêtre contient entre 60 et 90% de surface agricole (en vert).

Il est important de rappeler que ces zones correspondent aux pixels qui pourront potentiellement être échantillonnés par la suite, c'est-à-dire aux pixels centraux des carrés de n kilomètres de côté qui nous intéressent. Ainsi une zone peut être très morcelée voire de petite taille (comme les petites taches vertes isolées en bas à droite de l'image centrale de la figure 13) et néanmoins servir à l'échantillonnage. Chaque pixel renseigne sur une étendue correspondant à la fenêtre utilisée et prend en compte ses voisins dans cette étendue.

3) Calcul des rasters basés sur la taille moyenne de *patch*.

Dans un premier temps nous avons calculé le raster global basé sur la taille moyenne de *patch* à l'intérieur de la fenêtre mobile. Les rasters basés sur les fenêtres de tailles de 0,25, 1 et 3 kilomètres sont données dans la figure 14.

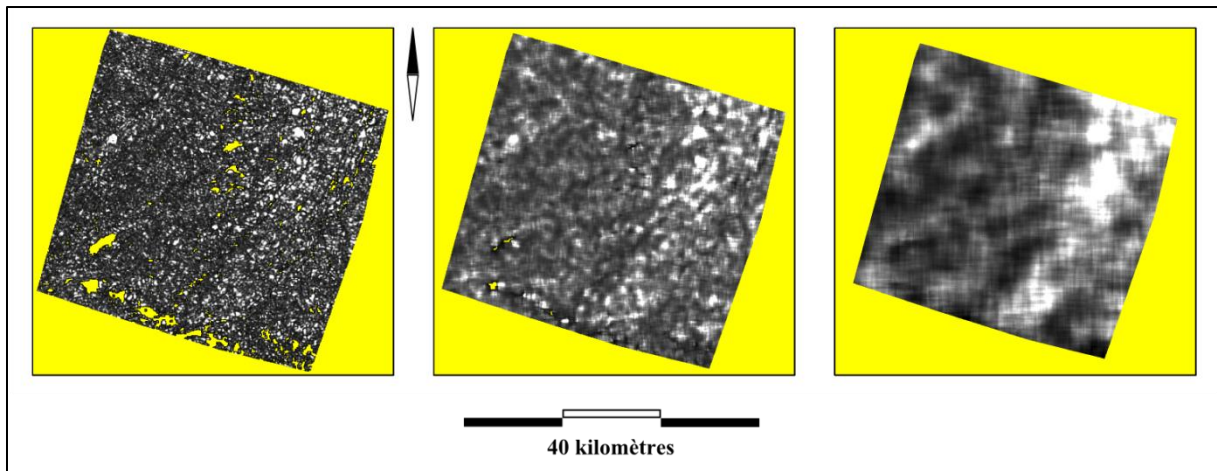


Figure 14 : Rasters basés sur la taille moyenne de *patch* calculée avec des fenêtres mobiles de 0,25, 1 et 3 kilomètres. Plus le raster est clair plus la taille moyenne de *patch* est élevée et plus l'hétérogénéité de configuration est faible. Les zones en jaune retrouvées dans l'emprise de la zone étudiée correspondent à des localités pour lesquelles la fenêtre mobile ne contenait pas de pixels renvoyant à une modalité d'occupation du sol de type agricole (données manquantes ou hors champ).

Dans un second temps, nous avons appliqué le masque calculé précédemment afin de ne retenir que les pixels situés dans les zones comportant entre 60 et 90% de surface agricole. Ces résultats apparaissent sur la figure 15 présentant les rasters construits à partir des fenêtres de 0,25, 1 et 3 kilomètres.

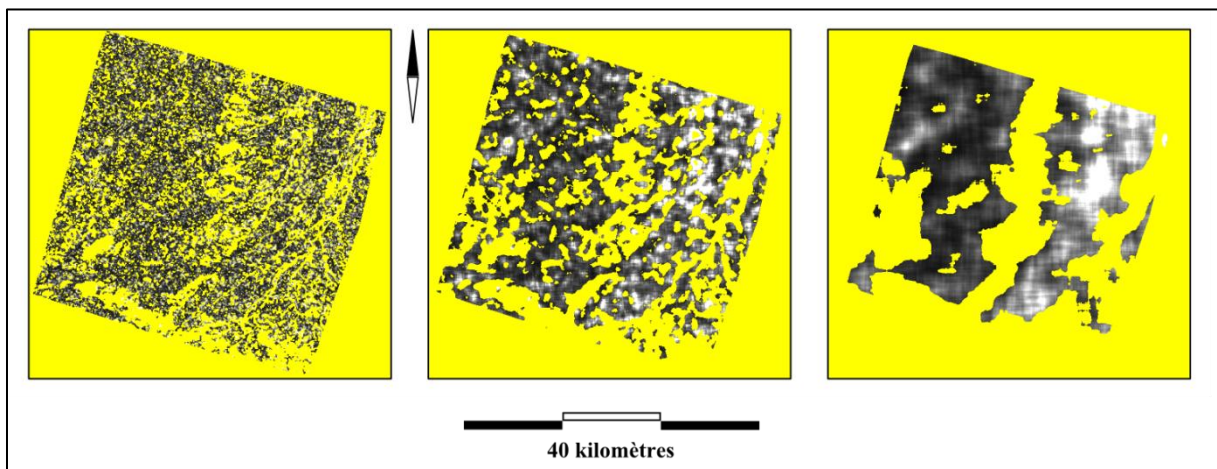


Figure 15 : Rasters basés sur la taille moyenne de *patch* retenant les pixels centraux des fenêtres mobiles de 0,25, 1 et 3 kilomètres pour lesquelles la fenêtre contient entre 60 et 90% de surface agricole. Plus le raster est clair plus la taille moyenne de *patch* est élevée et plus l'hétérogénéité de configuration est faible. Les zones en jaune retrouvées dans l'emprise de la zone étudiée correspondent à des localités pour lesquelles la fenêtre mobile ne contenait pas de pixels renvoyant à une modalité d'occupation du sol de type agricole (données manquantes ou hors champ).

4) Calcul des rasters basés sur l'indice de diversité de Shannon.

De la même manière, nous avons calculé le raster global basé sur l'indice de diversité de Shannon pour les 8 tailles de fenêtre retenues précédemment dont trois cas apparaissent sur la figure 16.

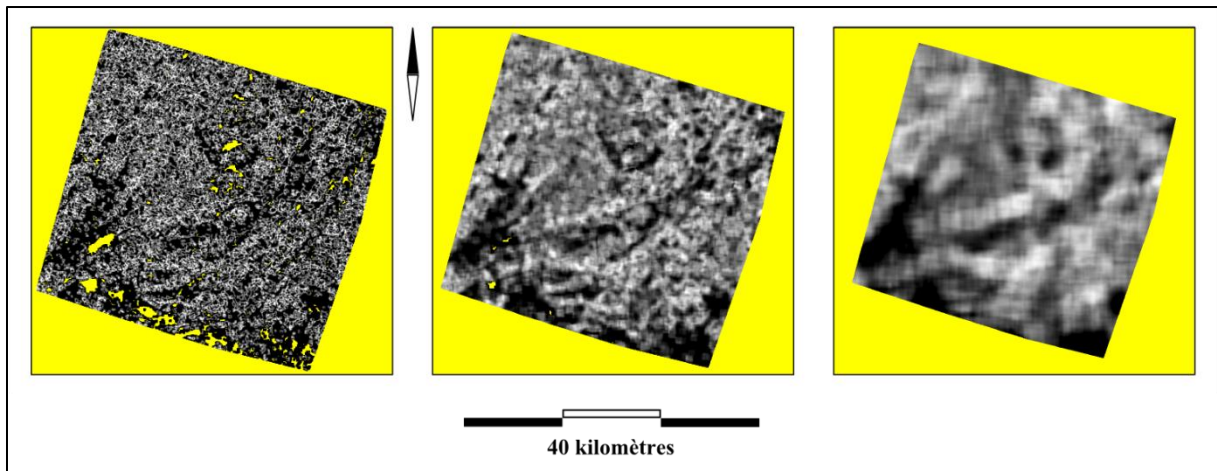


Figure 16 : Rasters basés sur l'indice de diversité de Shannon calculé avec des fenêtres mobiles de 0,25, 1 et 3 kilomètres. Plus le raster est clair plus l'indice de Shannon est élevé et plus l'hétérogénéité de composition est forte. Les zones en jaune retrouvées dans l'emprise de la zone étudiée correspondent à des localités pour lesquelles la fenêtre mobile ne contenait pas de pixels renvoyant à une modalité d'occupation du sol de type agricole (données manquantes ou hors champ).

Puis nous avons appliqué sur cette dernière série d'images le masque retenant les pixels situés dans les zones comprenant entre 60 et 90% de surface agricole. Le résultat apparaît dans la figure 17.

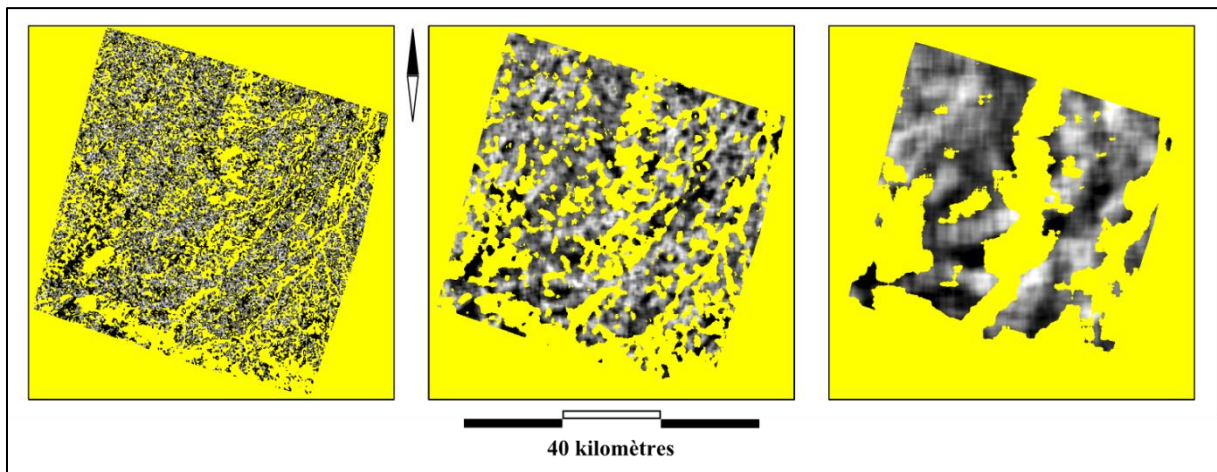


Figure 17 : Rasters basés sur l'indice de diversité de Shannon retenant les pixels centraux des fenêtres mobiles de 0,25, 1 et 3 kilomètres pour lesquelles la fenêtre contient entre 60 et 90% de surface agricole. Plus le raster est clair plus l'indice de Shannon est élevé et plus l'hétérogénéité de composition est forte. Les zones en jaune retrouvées dans l'emprise de la zone étudiée correspondent à des localités pour lesquelles la fenêtre mobile ne contenait pas de pixels renvoyant à une modalité d'occupation du sol de type agricole (données manquantes ou hors champ).

5) Estimation des rasters à conserver : choix de la taille de la fenêtre mobile.

Le but de cette étape dans le protocole initial mis en place par Pasher *et al.* est de maximiser la variation de l'hétérogénéité des paysages en restreignant le nombre d'échelles d'analyse. En effet, toutes les tailles de fenêtre choisies initialement ne sont pas pertinentes, et seul un nombre limité d'entre elles pourrait apporter de l'information. Les fenêtres mobiles

FRAGSTATS se comportant comme un filtre de lissage, à partir d'une certaine taille de fenêtre la variabilité entre points tend à se stabiliser autour d'une valeur moyenne. Il sera donc inutile d'aller au-delà d'une certaine taille de fenêtre. Par ailleurs, pour les fenêtres de trop petites tailles, des artefacts peuvent apparaître. Par exemple, le nombre maximal de parcelles dans un carré d'une taille donnée est limité par la taille moyenne des parcelles.

Dans FARMLAND, une taille unique de fenêtre a été choisie *a priori* pour des quadrats (1 km de côté) afin de standardiser les comparaisons entre pays. Je me suis donc assuré que cette taille de fenêtre avait une certaine pertinence statistique. Nous avons également proposé d'autres tailles de fenêtre en cohérence avec le protocole de Pasher *et al.*.

Afin d'estimer la limite supérieure de taille de fenêtre mobiles à partir de laquelle nous ne gagnons plus d'information, Pasher *et al.* suggèrent de calculer, pour chaque métrique et pour chaque étendue de paysage, le coefficient de variation pour l'ensemble des pixels situés dans les zones comprenant de 60 à 90% de surface agricole. Ce paramètre statistique permet de comparer les différentes gammes de variation pour chaque taille de fenêtre en s'affranchissant des valeurs moyennes. La représentation graphique de l'évolution des valeurs de ce coefficient de variation en fonction de la taille de la fenêtre pour la taille moyenne de tache de végétation et pour l'indice de Shannon apparaît sur la figure 18. Cette figure montre à partir de quelle taille de fenêtre la dispersion des valeurs se stabilise et donc cesse d'apporter de l'information (approximativement pour des tailles de fenêtre mobile supérieures à 2 ou 3 kilomètres).

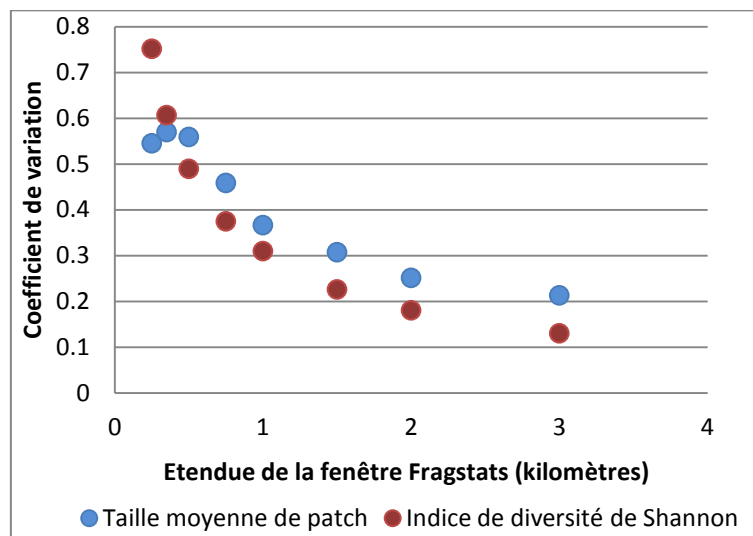


Figure 18 : Evolution du coefficient de variation de la taille moyenne de patch et de l'indice de Shannon en fonction des 8 tailles de fenêtre mobiles testées.

Il peut être important de noter que les dynamiques de ces évolutions ne sont pas exactement les mêmes pour l'indice de Shannon et la taille moyenne de patch. Si la première courbe suit une hyperbole décroissante, la deuxième présente pour ses 3 premières valeurs un plateau avant de décroître. Ce plateau correspond probablement à un artefact. La variation de la valeur de la taille de patch ne doit que peu augmenter localement pour les tailles de fenêtre faibles mais augmenter subitement après un seuil (quand un grand champ de maïs vient se perdre au milieu de petites prairies par exemple). Au contraire, la variation de la diversité doit

de son côté augmenter beaucoup plus vite. On peut émettre l'hypothèse que deux champs côte à côte ont plus de chance d'avoir la même taille qu'une même occupation du sol.

Afin de guider le choix de l'étendue à partir de laquelle le gain d'information est minime, Pasher *et al.* proposent de calculer la corrélation entre différentes tailles de fenêtres successives pour chaque métrique. Si pour une taille intermédiaire la corrélation est sensiblement la même que pour la taille forte, le choix de la première valeur devrait permettre d'arriver au même résultat de manière plus économique (et notamment économique en temps de calcul). Cette méthode permettrait ainsi de déterminer la plus grande taille de fenêtre à sélectionner. En suivant ce protocole, Pasher *et al.*, en étaient arrivés à sélectionner trois tailles de fenêtres mobiles sur leur terrain d'étude canadien respectivement 1, 2 et 3 kilomètres de côté. Cette approche nous a toutefois semblé poser certaines difficultés.

Malgré le calcul d'une statistique, ce choix n'est pas très différent d'un critère purement visuel. En effet, Pasher *et al.* considèrent les corrélations de chacune des deux métriques pour des fenêtres successives c'est-à-dire de 5 et 4, puis 4 et 3 puis 3 et 2 kilomètres. De cette manière, ces chercheurs trouvent que les corrélations restent fortes dans tous les cas mais qu'un léger saut est observé entre les corrélations des métriques basées sur des fenêtres de 3 et 2 kilomètres. Sur cette base, ils ont donc choisi de conserver la valeur de 3 kilomètres comme taille maximale de fenêtre. Ce choix ne nous semble pas plus étayé qu'un simple critère visuel sur les courbes car aucun test inférentiel ne vient le confirmer. Même si chaque corrélation est testée, le test porte sur la significativité de la corrélation et pas sur l'importance du saut entre deux fenêtres successives. De plus, on peut s'attendre à ce que la corrélation entre deux tailles successives de fenêtre soit plutôt forte (ce que nous pourrions voir plus loin dans les tableaux 6 à 9).

Sur la base de ces remarques, nous avons proposé une alternative en calculant la corrélation pour chaque métrique entre une taille de fenêtre de départ et toutes les tailles suivantes. Cela revient non pas à chercher si la corrélation reste forte, mais à rechercher quand elle cesse d'être significative.

Pour cela, une centaine de points ont été tirés aléatoirement dans le paysage à l'intérieur de l'emprise commune à tous les masques calculés. Ces points ont par la suite reçu dans leur table attributaire la valeur des comptes numériques des pixels des rasters calculés à partir des deux métriques FRAGSTATS. Cela nous a permis d'obtenir 8 lots de points pour la taille moyenne de tache et 8 lots pour l'indice de Shannon. Dans la mesure où les données ne sont pas distribuées normalement (ce qui a été vérifié graphiquement par des représentations des distributions et de manière inférentielle par des tests de Shapiro-Wilks) nous avons privilégié l'usage du coefficient de corrélation non-paramétrique de Spearman plutôt que son homologue paramétrique. Etant donné que des comparaisons multiples étaient en jeu, une correction de seuil de type Bonferroni a été utilisée pour le calcul des probabilités. Les résultats sont présentés dans les tableaux 5 et 6 pour les tailles moyennes de tache et 7 et 8 pour les indices de Shannon (valeurs des coefficients de corrélation et probabilités respectivement dans chaque cas). Dans la mesure où nos échantillons comportaient un nombre relativement élevé d'individus, et que les calculs de corrélation sont sensibles à la taille de

l'échantillon (ce qui est directement visible dans une table statistique de Spearman), nous avons utilisé un seuil de significativité sévère à 1 pour 1000 pour éviter des faux-positifs.

rho Spearman	Area 250	Area 350	Area 500	Area 750	Area 1000	Area 1500	Area 2000	Area 3000
area250	1							
area350	0.67	1						
area500	0.4	0.64	1					
area750	0.36	0.51	0.64	1				
area1000	0.29	0.41	0.5	0.78	1			
area1500	0.28	0.37	0.49	0.62	0.77	1		
area2000	0.24	0.36	0.39	0.52	0.69	0.86	1	
area3000	0.18	0.32	0.37	0.43	0.55	0.66	0.85	1

Tableau 5: Matrices des corrélations de Spearman décrivant les relations entre les comptes numériques des pixels des rasters basés sur la taille moyenne de patch pour les différentes tailles de fenêtre mobile. La valeur surlignée indique le seuil à partir duquel la corrélation cesse d'être significative au seuil choisi.

p	Area 250	Area 350	Area 500	Area 750	Area 1000	Area 1500	Area 2000	Area 3000
area250								
area350	0							
area500	0.0003	0						
area750	0.0016	0	0					
area1000	0.0148	0.0002	0	0				
area1500	0.0148	0.0013	0	0	0			
area2000	0.0312	0.0016	0.0005	0	0	0		
area3000	0.071	0.0061	0.0014	0.0001	0	0	0	

Tableau 6: Matrices des valeurs p relatives aux corrélations de Spearman entre les comptes numériques des pixels des rasters basés sur la taille moyenne de patch pour les différentes tailles de fenêtre mobile. La valeur surlignée indique le seuil à partir duquel la corrélation cesse d'être significative au seuil choisi.

rho Spearman	SHDI 250	SHDI 350	SHDI 500	SHDI 750	SHDI 1000	SHDI 1500	SHDI 2000	SHDI 3000
SHDI250	1							
SHDI350	0,93	1						
SHDI500	0,76	0,9	1					
SHDI750	0,57	0,69	0,87	1				
SHDI1000	0,44	0,53	0,69	0,9	1			
SHDI1500	0,32	0,39	0,52	0,7	0,88	1		
SHDI2000	0,18	0,26	0,39	0,53	0,7	0,9	1	
SHDI3000	0,06	0,1	0,18	0,25	0,42	0,66	0,83	1

Tableau 7: Matrices des corrélations de Spearman décrivant les relations entre les comptes numériques des pixels des rasters basés sur le calcul de l'indice de diversité de Shannon pour les différentes tailles de fenêtre mobile. La valeur surlignée indique le seuil à partir duquel la corrélation cesse d'être significative au seuil choisi.

p	SHDI 250	SHDI 350	SHDI 500	SHDI 750	SHDI 1000	SHDI 1500	SHDI 2000	SHDI 3000
SHDI250								
SHDI350	0							
SHDI500	0	0						
SHDI750	0	0	0					
SHDI1000	0	0	0	0				
SHDI1500	0.0089	0.0005	0	0	0			
SHDI2000	0.2743	0.0475	0.0005	0	0	0		
SHDI3000	0.6911	0.6911	0.2743	0.0578	0.0001	0	0	

Tableau 8: Matrices des valeurs p relatives aux corrélations de Spearman entre les comptes numériques des pixels des rasters basés sur le calcul de l'indice de diversité de Shannon pour les différentes tailles de fenêtre mobile. La valeur surlignée indique le seuil à partir duquel la corrélation cesse d'être significative au seuil choisi.

Il reste à déterminer la borne inférieure de la plage de valeurs. Cette limite va notamment être posée par la structure et la configuration des parcelles. Comme cela a pu être évoqué précédemment, la taille moyenne des parcelles va contraindre leur nombre maximum pour une taille de fenêtre donnée. Cela est particulièrement évident si l'on effectue un échantillonnage aléatoire d'un grand nombre de points dans le masque des zones ayant 60 à 90% de surface agricole, que l'on extrait l'information des deux indices d'hétérogénéité et que l'on réalise le nuage de points croisant ces deux indices (figure 19). La structuration de l'espace agricole crée un artefact facilement décelable pour la plus petite taille de fenêtre avec des paliers nets dans le nuage de points. Cet artefact reste visible pour une taille de fenêtre de 500 mètres de côté et s'estompe nettement pour une fenêtre de 1 kilomètre carré. Cette taille de fenêtre de 1 kilomètre de côté pourrait donc être considérée dans notre cas comme la taille minimale à choisir pour une fenêtre mobile. Pour cette taille minimale, les tableaux 5 à 8 nous informent que les tailles de fenêtre de 1,5 à 3 kilomètres peuvent être choisies. La probabilité associée à la corrélation entre les valeurs de l'indice de Shannon pour une fenêtre de 1 kilomètre de côté et une fenêtre de 3 kilomètres de côté commence toutefois à augmenter et pourrait ne plus être significative au-delà.

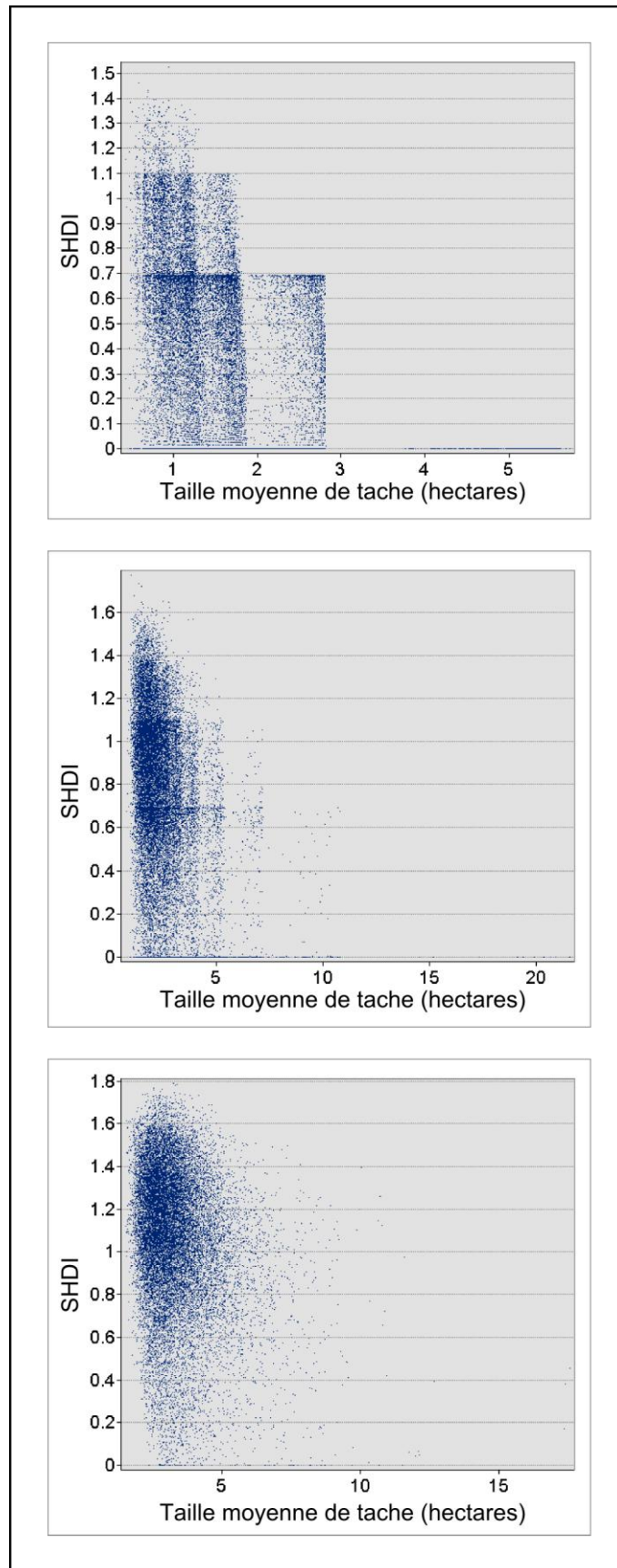


Figure 19: Nuages de points réalisés à partir d'un échantillonnage aléatoire de 20000 points dans les masques comportant entre 60 et 90 % de surface agricoles et renseignés par les rasters des deux indices d'hétérogénéité pour les fenêtres mobiles de 0,25 ; 0,5 et 1 kilomètre de côté (respectivement de haut en bas).

Il apparaît à l'étude des tableaux 5 à 8 précédents et des figures 18 et 19 que la taille minimale des fenêtres mobiles doit se situer autour de 1 km de côté et que le maximum se situe probablement peu au-delà de 3 kilomètres de côté. Toutefois avec cette dernière taille de fenêtre, même si la corrélation reste significative, sa valeur tend à diminuer nettement par rapport aux fenêtres plus petites. Si d'autres échelles d'analyses devaient être testées dans le cadre du projet FARMLAND, des fenêtres de 1,5 ou 2 kilomètres seraient probablement plus intéressantes que des fenêtres de 3 kilomètres de côté.

Le choix actuel d'une fenêtre de 1 kilomètre de côté dans FARMLAND semble donc pertinent compte tenu des caractéristiques des paysages des Coteaux de Gascogne.

6) Estimation de la distance minimale entre deux fenêtres mobiles.

Le but de cette sous-étape va être d'éviter l'autocorrélation spatiale entre les paysages pour chaque étendue conservée. En effet, à terme, les carrés d'échantillonnage devront être à une certaine distance les uns des autres pour éviter cette autocorrélation. A cette fin, Pasher *et al.* proposent de calculer un indice d'autocorrélation, le I de Moran (c.f. annexes), pour chaque métrique, pour chaque étendue sélectionnée précédemment et pour des pas de distances allant de 30 m (résolution du pixel dans leur cas) à 5 km. Compte tenu de nos données et du temps nécessaire à calculer les cartes de I de Moran sous ENVI v4.7 et les statistiques qui en découlent, nous avons restreint ce pas à l'intervalle [0,01-2,5 kilomètres] (rappel : la résolution du pixel dans notre cas est de 10 mètres). Le point d'inflexion des courbes renseigne sur la distance minimale à mettre entre les centres des carrés d'échantillonnage. La figure 21 montre l'évolution de ces courbes pour les tailles de fenêtre mobiles de taille 1 ; 1,5 et 2 kilomètres.

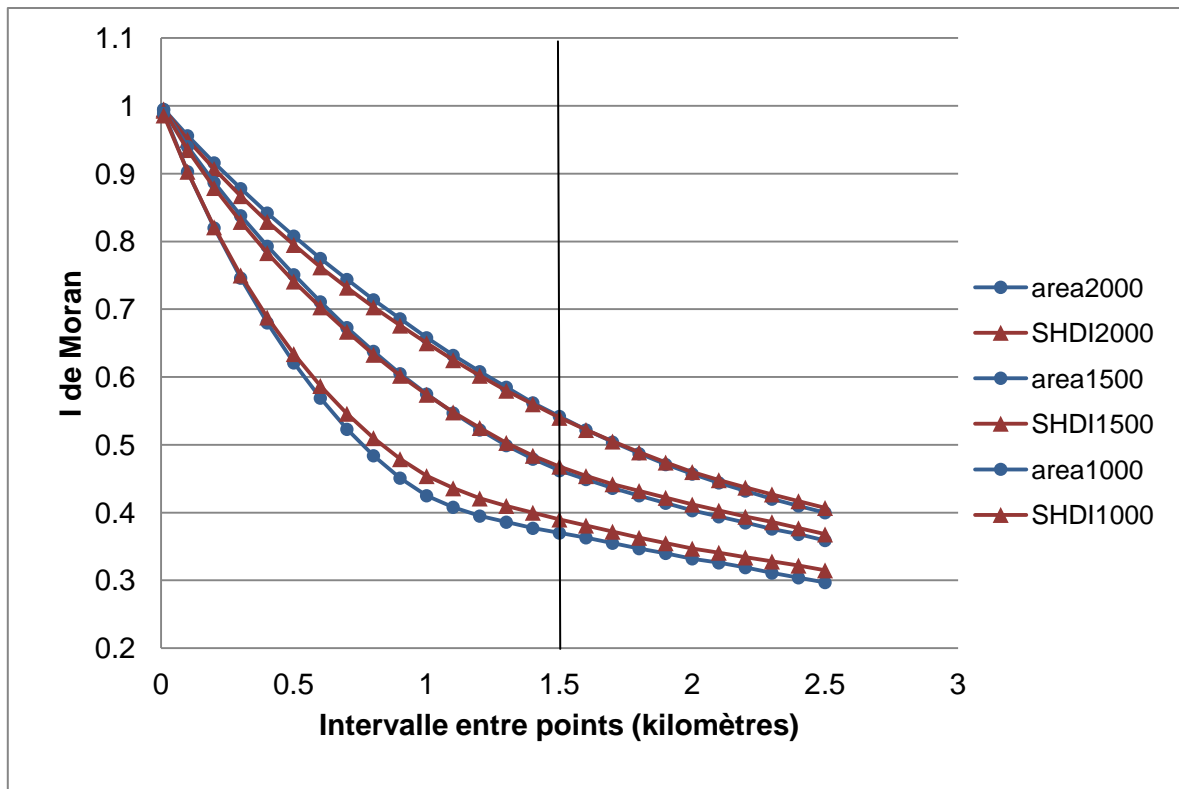


Figure 21 : Evolution de l'indice de Moran pour les rasters basés sur la taille moyenne de tache pour trois tailles de fenêtre mobile en fonction de l'intervalle choisi (en kilomètres). Une distance minimale de 1500m entre deux centres de fenêtre a été choisie.

Il apparaît très naturellement que plus la taille de la fenêtre mobile augmente plus il faut s'éloigner d'un pixel pour que les données ne soient pas corrélées spatialement. Pasher *et al.*, choisissent une distance de 3,5 km entre les centres des différents points d'échantillonnage et une taille maximale de fenêtre mobile de 3 km. Après analyse détaillée de cette configuration spatiale d'échantillonnage, cette valeur nous semble être insuffisante. Si elle permet de limiter le recouvrement des quadrats, elle ne l'interdit pas et deux zones d'échantillonnage de la biodiversité pourraient se recouvrir par un de leurs sommets. En effet, la diagonale d'un carré de 3 kilomètres de côté mesure approximativement 4,25 kilomètres. Autrement dit, une distance entre barycentres de deux quadrats de 3,5 kilomètres autoriserait bien un recouvrement. Nous avons fait remonter ce point technique à Pasher *et al.* pour qu'ils en tiennent compte dans leur étude.

Aussi compte-tenu de ce point, dans notre cas, nous proposons de choisir une distance entre centres de quadrats d'échantillonnage de 1,5 km soit une distance minimale d'un peu plus de 100 mètres entre les sommets de deux quadrats pour des carrés de 1 kilomètre de côté (pour lesquels la diagonale mesure approximativement 1,4 kilomètre).

7) Discussion sur le choix des tailles de fenêtre mobile à conserver et sur l'estimation de la distance minimale entre centre des quadrats d'échantillonnage.

Cette longue étape nous a permis de confirmer le bien-fondé du choix de fenêtre de 1 kilomètre de côté fait initialement à l'écriture du projet FARMLAND à dire d'expert. Cette mesure est pertinente compte tenu des caractéristiques de nos paysages et permet de conserver une part importante de la variabilité qui aurait été perdue avec des fenêtres plus grandes (au-delà de 3 kilomètres).

Sur la base de discussion avec Lenore Fahrig, nous avons proposé une méthode complémentaire pour le choix des tailles de fenêtres. Cette méthode prend en compte la variabilité des données qui permet d'avoir un contact très visuel des données, mais aussi une approche plus inférentielle. Enfin cette méthode prend en compte les artefacts éventuels pouvant apparaître avec de faible taille de fenêtre.

Sur la base du calcul du I de Moran nous avons pu déterminer une distance minimale entre quadrats d'échantillonnage de la biodiversité. Cette information sera essentielle pour l'étape suivante de sélection de ces quadrats. Il est important de noter que cette autocorrélation est toujours restée significative, même une fois que l'indice tend vers une valeur plancher. Le choix d'une fenêtre de 1 kilomètre de côté présentant des valeurs d'autocorrélation plus faible que les fenêtres de taille supérieure contribue à limiter les biais occasionnés. Selon Lenore Fahrig (communication personnelle), l'idéal serait de pouvoir étudier le I de Moran des résidus de la relation entre la biodiversité et les indices d'hétérogénéité, tout le problème étant que l'on ne peut connaître cela qu'après avoir effectué les relevés de biodiversité (et donc une fois que l'on n'a plus besoin de l'information sur l'autocorrélation)... Une perspective d'amélioration de la phase d'estimation de l'autocorrélation nous a été suggérée par Michel Goulard, géostatisticien de l'UMR DYNAFOR. Celle-ci pourrait consister à calculer le I de Moran non pas sur les valeurs réelles des rasters de départ, mais sur l'écart de celles-ci à un modèle de référence (un filtre médian de l'image de départ par exemple). Cela pourrait permettre de s'affranchir de l'effet de structures inhérentes au paysage venant biaiser les calculs d'autocorrélation. Nous n'avons toutefois pas eu le temps de mettre cette alternative en pratique et celle-ci reste à explorer.

IX. Etape 5 : Echantillonnage des points de relevés des suivis faunistiques et floristiques.

A partir de l'étape précédente, nous avons pu obtenir les éléments nous permettant d'envisager la sélection des quadrats d'échantillonnage de la biodiversité. Nous disposons de rasters des deux indices d'hétérogénéité de composition et de configuration. Nous avons confirmé la pertinence de l'échelle d'analyse choisie et nous disposons d'une distance minimale entre les bords de ces quadrats. Comme cela a été évoqué précédemment, FRAGSTATS v3.3 calcule les rasters basés sur les indices d'hétérogénéité du paysage par le biais d'une fenêtre mobile se comportant comme un filtre de lissage. La valeur de chaque pixel est relative aux pixels voisins compris dans la fenêtre et la sélection de nos quadrats pourra se faire sur la base de la valeur de leur pixel central. Il conviendra que ce pixel central

correspondre à l'une des combinaisons souhaitées de valeurs d'indices d'hétérogénéité du paysage agricole. Afin de « forcer le trait » et de se placer dans des conditions expérimentales optimales, Fahrig *et al.* (2010) suggèrent de réaliser une discrétisation des valeurs de ces indices et de ne sélectionner que les quatre combinaisons des modalités extrêmes de chacun d'entre eux (faible-faible, faible-fort, fort-faible et fort-fort). Une faible hétérogénéité de configuration correspondra à des tailles moyennes de patch forte. Afin de pouvoir dissocier les effets de ces deux dimensions de l'hétérogénéité, il faudra s'assurer de leur indépendance statistique.

1) Détermination des plages de valeur correspondant aux quatre combinaisons de modalités d'hétérogénéité et recodage des rasters d'hétérogénéité.

Afin de déterminer les valeurs extrêmes de la gamme des valeurs de chaque indice d'hétérogénéité, il convient d'avoir un aperçu de l'étendue de ces valeurs. Grâce à l'extension Hawth's Tools (Hawth's Analysis Tools, 2004) installable dans ArcGis, il est aisé de réaliser un plan d'échantillonnage (20000 points) à l'intérieur du masque retenant entre 60 et 90% de surface agricole avec la fonction « Generate Random Points ». Par la suite, il est possible d'extraire l'information des différents rasters basés soit sur la taille moyenne de tache soit sur l'indice de diversité de Shannon et d'exporter le résultat dans un tableur pour en faire une représentation graphique sous la forme d'un nuage de points avec en abscisse les valeurs de taille moyenne de tache et en ordonnée les valeurs de l'indice de Shannon (comme cela a pu être réalisé dans la figure 19 lors de l'étape 5). Fahrig *et al.* (2010) recommandent toutefois de retirer les valeurs extrêmes de la distribution de chacune de ces variables (2,5% en tête et queue de chaque distribution) afin d'éliminer les événements trop rares ou aberrants.

Ces valeurs ont par la suite été recodées en déciles (intervalles égaux) afin de ne conserver que les combinaisons de modalités les plus extrêmes tout en garantissant un nombre suffisant de points pour notre échantillonnage final et une distribution homogène dans l'espace. La figure 22 confirme la très faible relation qu'il y a entre nos deux variables. Un des objectifs de ce travail était de déterminer une cinquantaine de points d'échantillonnage de quatre types correspondant à des combinaisons de modalités d'indice d'hétérogénéité de composition (Comp.) et de configuration (Conf.): Faible-Conf/Faible-Comp, Forte-Conf /Faible-Comp, Faible-Conf / Forte-Comp, et Forte-Conf /Forte-Comp. Compte tenu de la dispersion des valeurs dans le nuage de points, les plages de valeurs contenant un nombre de points répartis de manière homogène dans l'aire d'étude ne vont pas être de même étendues pour chaque combinaison. Nous avons donc augmenté la taille de ces plages de valeurs par pas de décile de manière à obtenir des sous-nuages de points qui soient homogènes tout en veillant à ce qu'ils correspondent à des valeurs les plus extrêmes possible sur chaque axe. Afin d'évaluer cette homogénéité, nous avons calculé la corrélation non paramétrique de Spearman entre nos deux indices d'hétérogénéité. Compte tenu de la taille importante de nos effectifs, nous avons appliqué un seuil de significativité sévère de 1%. Les combinaisons retenues vont ainsi correspondre aux plages de déciles 1-3/1-3, 8-10/1-4, 1/10 et 7-10/7-10 (la première plage de valeurs étant relative à la taille moyenne de tache et la seconde à l'indice de diversité de Shannon). Le fait de réduire la plage des valeurs des combinaisons Faible-Conf./Faible-Comp. et Forte-Conf./Forte-Comp. conduit à avoir des effectifs plus équilibrés,

mais des valeurs du rho de Spearman plus élevées et des probabilités associées plus faibles. Nous avons donc préféré conserver les plages de valeurs précédemment citées.

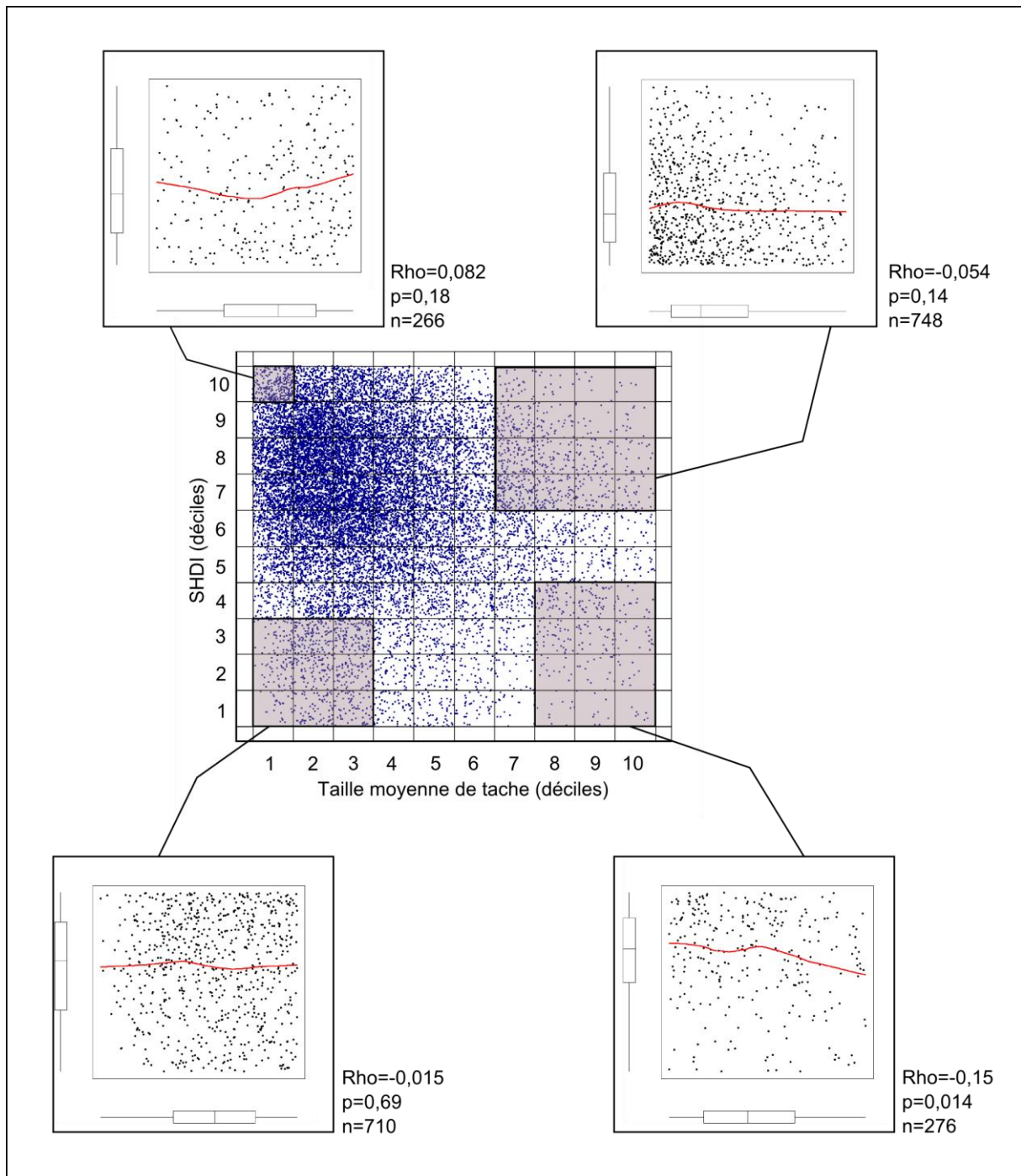


Figure 22 : Nouvel échantillonnage réalisé sur le nuage de points afin d'extraire 4 lots de points correspondants aux 4 combinaisons d'indice d'hétérogénéité de composition et de configuration (a, b, c, et d de la figure 2). Les valeurs extrêmes de la distribution de chacune des variables ont été éliminées. Pour chaque nouveau sous-groupe le coefficient de corrélation de Spearman entre les deux indices a été calculé afin de s'assurer de l'indépendance des deux variables dans chaque sous-groupe et une courbe de lissage non-paramétrique décrit la relation entre les deux indices d'hétérogénéité.

Il est alors possible de recoder les rasters de métriques issus de Fragstats (après application du masque retenant les zones comprenant entre 60 et 90% de surface agricole) importés dans ArcGis en reclassant les points selon les plages de valeurs décrites par les déciles et le croisement de ces deux rasters a permis de faire émerger les différentes combinaisons possibles (figure 23). Les autres occupations du sol ont été recodées en données manquantes.

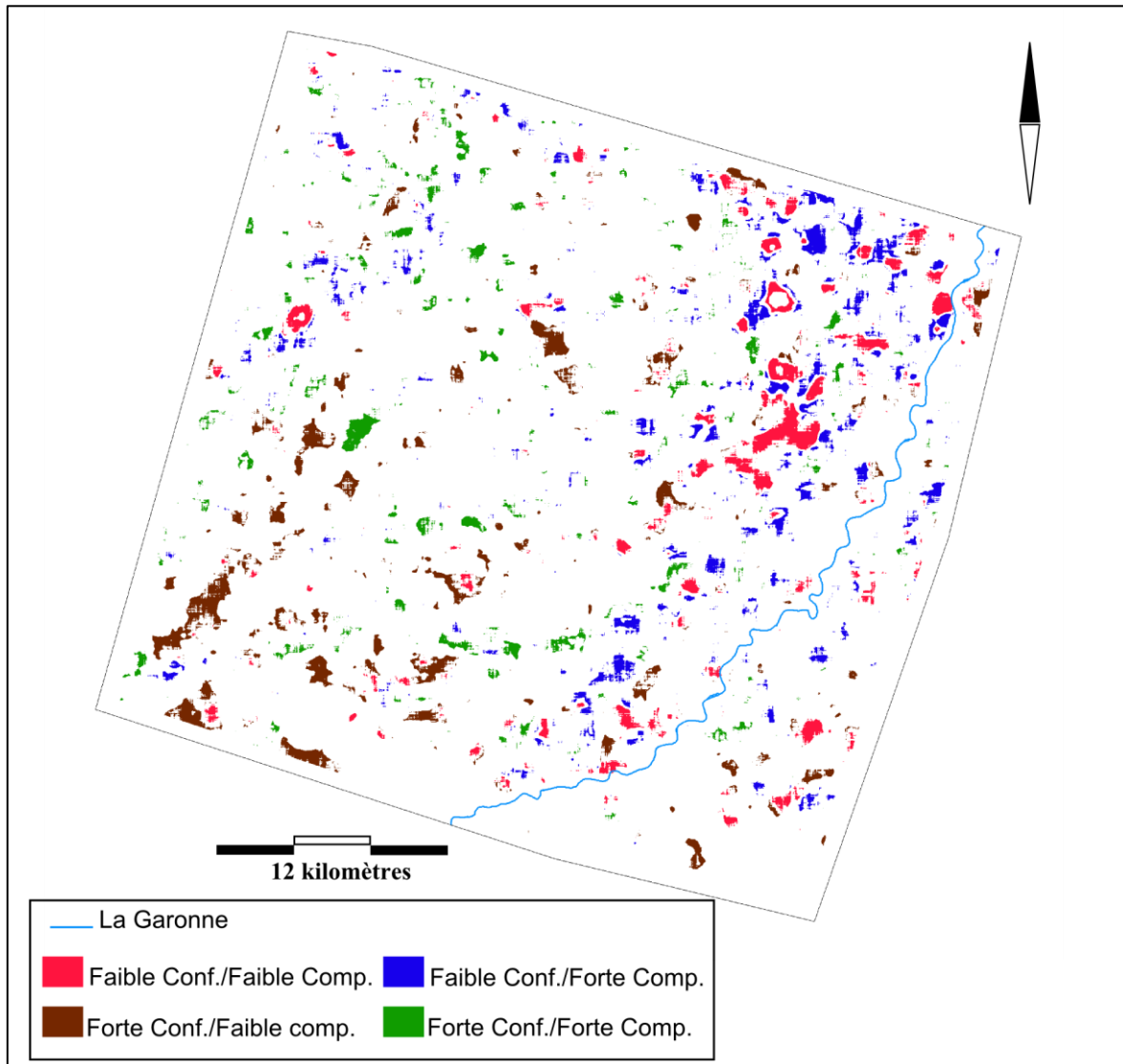


Figure 23 : Localisation géographique des pixels correspondants aux quatre combinaisons d'indices d'hétérogénéité de configuration (Conf.) et de composition (Comp.) pour les zones comprenant entre 60 et 90% de surface agricole.

2) Echantillonnage des barycentres des quadrats.

Grace à l'extension Hawth's Tools, il est aisé de projeter un semis de points distants au minimum de 1500 mètres les uns des autres dans l'emprise du raster obtenu dernièrement, à l'exclusion des pixels recodés en données manquantes. Ces points peuvent alors être renseignés par le compte numérique des pixels du raster qui correspondent à nos différentes

combinaisons de modalités d'hétérogénéité. Un champ supplémentaire a été rajouté à la table attributaire et contenant un nombre réel compris entre 0 et 1 tiré aléatoirement. Ce champ servira à sélectionner au hasard des quadrats sur l'aire d'étude et à éviter une agrégation de points dans une partie de l'image. A l'aide de ces deux derniers champs (compte numérique renseignant sur l'occupation du sol et nombre aléatoire), il est possible de réaliser la sélection de 12 points par type d'occupation du sol, se situant à 1500 mètres au moins les uns des autres et répartis de manière homogène sur le territoire (sans agrégation d'une modalité à un endroit donné). Si la répartition spatiale est malgré tout jugée visuellement trop déséquilibrée, il est tout à fait possible d'y remédier en remplaçant un point par un autre manuellement. La figure 24 présente un exemple de la distribution spatiale de ces points pour chaque combinaison de modalités d'hétérogénéité.

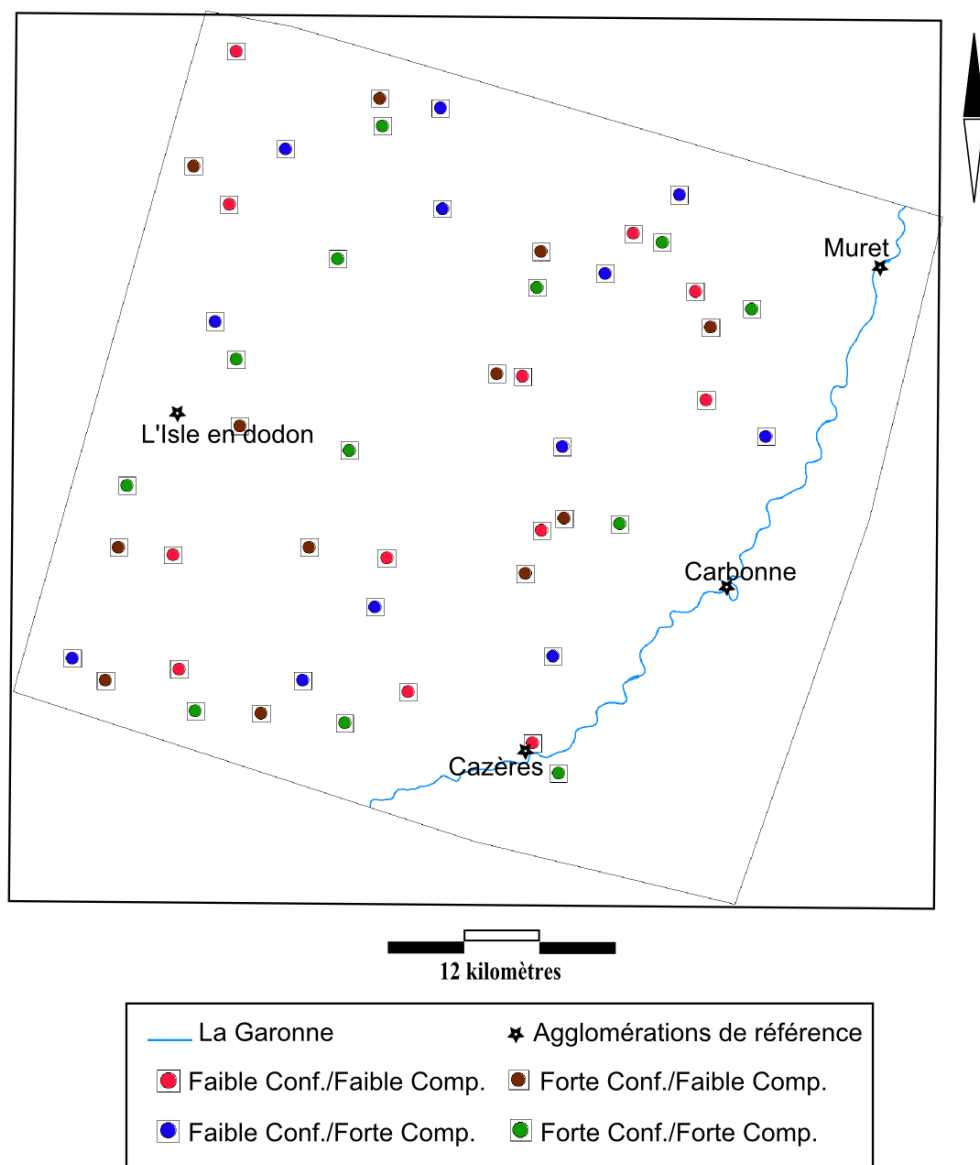


Figure 24 : Distribution spatiale des quadrats d'échantillonnage correspondants à chacune des combinaisons de modalités d'indices d'hétérogénéité de configuration (Conf.) et de composition (Comp.). Chaque point représente le barycentre d'un quadrat et le carré dans lequel il s'inscrit représente son emprise.

3) Vérification terrain.

Suite à cet échantillonnage, nous avons effectué une vérification de terrain en allant contrôler *in situ* la conformité de notre sélection avec le cahier des charges initial. En vue de cette vérification nous avons réalisé un atlas cartographique des différents quadrats d'études comprenant l'emprise de chaque quadrat de 1 kilomètre de côté sur un fond de carte extrait de la base de donnée ORTHO 2006 de l'Institut Géographique National et un quadrillage géographique avec un maillage de 100 mètres. Cet atlas a été réalisé en utilisant l'outil « Créer un atlas cartographique » téléchargeable sur la suite du support technique ESRI France⁸.

Ces supports cartographiques nous ont permis d'effectuer une première vérification au bureau pour nous assurer qu'un quadrat ne soit pas retenu dans une zone en construction ou dans une zone difficile d'accès. Dans un second temps, nous avons réalisé une sortie sur le terrain sur un quart des quadrats retenus (soit 12 sur 48 quadrats) avec une facilité de localisation grâce au quadrillage et à la photo aérienne couleur.

Un des grands intérêts de ce contact avec une réalité plus tangible que des données vues du ciel tient dans ce qu'il permet d'en relever les points d'achoppement. Comme cela a été évoqué lors de différentes étapes de notre cheminement, la détermination de l'occupation du sol n'a pas été parfaite dans la mesure où le pourcentage de pixels bien classé n'atteignait qu'un peu plus de 80%. Près du cinquième des pixels demeurait donc mal classé. Alors que les modalités d'occupation du sol de nature agricole sont difficilement évaluables sur des photographies aériennes, les zones bâties le sont beaucoup plus aisément.

Il serait alors facile d'identifier des erreurs de classification comme cela peut être le cas pour un quadrat visité, que nous avons retrouvé en plein lotissement, celui-ci ayant été identifié comme constitué essentiellement de prairies par confusion avec les pelouses et parcs attenants aux habitations. De la même manière, un quadrat a été positionné initialement à cheval sur l'enceinte d'une usine d'engins pyrotechniques. Cette fois encore, les pelouses entourant les bâtiments ont été identifiées comme des prairies. Au total, près d'une demi-douzaine de points d'échantillonnage nécessiteraient d'être repositionnés suite à ce type d'erreurs de classifications sur la seule base de comparaisons avec les photographies.

Il est toutefois important de souligner que cette vérification par photographie aérienne sera probablement améliorée par l'utilisation de la nouvelle BD ORTHO datant de 2010 disponible prochainement pour le département 31. En complément, l'utilisation de la base de données BD TOPO de l'Institut Géographique National et en particulier le thème des bâtiments (contenant la couche vectorielle du bâti) permettrait de lever en partie ces doutes car datant de 2008 elle permettrait de repérer des zones suspectes. Toutefois, seule une visite réelle du terrain apportera des certitudes sur le choix final des quadrats à inventorier car il y aura synchronisation temporelle complète entre état du paysage et relevé de biodiversité. Cela a notamment été le cas pour un des quadrats pour lequel un décalage important en termes de

⁸ <http://support.esrifrance.fr/index.asp?page=/outils/scripts/arcgis/arcmap/miseenpage/atlas/atlas.html>

surface bâtie a été observé avec la photographie aérienne : un lotissement a été construit depuis l'image satellitaire de 2009. Comme cela a été dit précédemment paragraphe IV-3, une piste d'amélioration de la classification de l'occupation du sol consisterait à effectuer des masques successifs sur le métafichier de départ afin d'éliminer des zones de la procédure de classification. Notamment, la vérification terrain via les orthophotographies a révélé qu'il pourrait être intéressant de s'affranchir des zones correspondant aux agglomérations. Cela pourrait être réalisé en utilisant une base de données les répertoriant (comme *CORINE Land Cover* pilotée par l'Agence Européenne de l'Environnement ou la base de données *TOPO* de l'Institut Géographique National). Bien que relativement anciennes (2006) et en dépit de niveaux d'échelles d'analyse différents (résolution minimale des objets de 25ha), l'utilisation des données *CORINE* aurait permis d'éviter une partie des erreurs de positionnement des quadrats liées à une mauvaise classification des surfaces bâties. Une des difficultés de ce travail consiste précisément à travailler à la fois à petite échelle sur des surfaces de grande étendue tout en ayant une certaine finesse d'analyse à plus grande échelle au niveau local. L'usage des données *CORINE* permettrait de délimiter les zones nous intéressant le moins, ce qui pourrait améliorer les performances de classification des zones de plus grand intérêt comme les surfaces agricoles. Toutefois, si cela permettait d'éliminer de relativement grandes agglomérations (tout est relatif !) comme les communes de Cazères ou Muret, cela ne permettrait pas de s'affranchir des petits villages et hameaux fréquemment rencontrés dans notre aire d'étude, trop petits pour figurer dans la typologie *CORINE* mais pouvant couvrir une part non négligeable d'un quadrat de 1 km².

Une alternative plus efficace pourrait consister à utiliser la couche relative au bâti de la base de données *TOPO* de l'Institut Géographique National (qui par ailleurs est plus récente que la base de données *CORINE*) et d'éliminer les zones comprenant plus d'une certaine densité d'habitations par une procédure de masquage. Un test a été effectué sur notre zone d'étude. Le calcul de densité d'habitation est facile à mettre en œuvre grâce à la boîte à outil de *Spatial Analyst* d'ArcGIS à partir des centroïdes des polygones représentant les bâtiments sur lesquels on applique la fonction « Densité simple ». Ces données pourraient alors intervenir à deux niveaux de la chaîne de traitement :

- Dans un premier temps, lors de la classification de l'occupation du sol, les zones impliquant plus d'une certaine densité d'agglomération sur le métafichier pourraient être masquées. Empiriquement l'utilisation d'un seuil de 100 habitations au kilomètre carré et le retrait des taches d'occupation du sol dont la surface est inférieure à un hectare semblent donner d'assez bons résultats en permettant d'éliminer bon nombre de groupes d'habitations et de hameaux tout en laissant de côté les innombrables habitations isolées (figure 25). Ces zones ont été appelées zones d'habitations. Faute de temps avant la date de remise du mémoire, nous n'avons pas pu refaire la procédure de classification après ce masquage, et celle-ci devra être réalisée par la suite.
- Dans un second temps, et plus particulièrement dans le cas présent où nous n'avons pas pu refaire la procédure de classification, il pourrait être intéressant de réaliser un masquage *a posteriori* des zones d'agglomération lors de la phase de sélection des

points d'échantillonnage. Nous définissons ces zones d'agglomération comme des zones intégrant plusieurs îlots de zones d'habitations non contigus, mais conservant une certaine proximité. La délimitation de ces zones permettra de tenir compte des pics locaux de croissance urbaine (typiquement une zone pavillonnaire en bordure d'une ville mais se détachant provisoirement de celle-ci). Pour cela, nous avons effectué un calcul de zone tampon autour des zones de densité d'habitations supérieure à 100 et dont la surface est supérieure à un hectare. La taille de l'agglomération semble devoir être considérée dans l'étendue de cette zone tampon pour mieux prendre en la croissance urbaine d'une année sur l'autre par exemple. Dans cet esprit, cette zone tampon n'a pas été appliquée de manière homogène, mais de manière proportionnelle à la surface d'agglomération détectée lors de l'étape précédente. Afin de ne pas exagérer les trop fortes valeurs, nous avons choisi de manière empirique une formule intégrant une fonction racine carré (50 fois la racine carrée de la surface des zones ayant une densité supérieure à 100 habitations au kilomètre carré et ayant une surface supérieure à 1 hectare). La fusion des différentes zones tampons recouvrantes permet de faire apparaître les contours des grandes zones d'agglomération de notre zone d'étude. Afin de ne pas être trop sévère en éliminant des zones correspondant à de petits hameaux, il nous a semblé judicieux d'appliquer un nouveau seuil. Empiriquement un seuil de 50 hectares nous a semblé intéressant pour éliminer la plupart des agglomérations (figure 25).

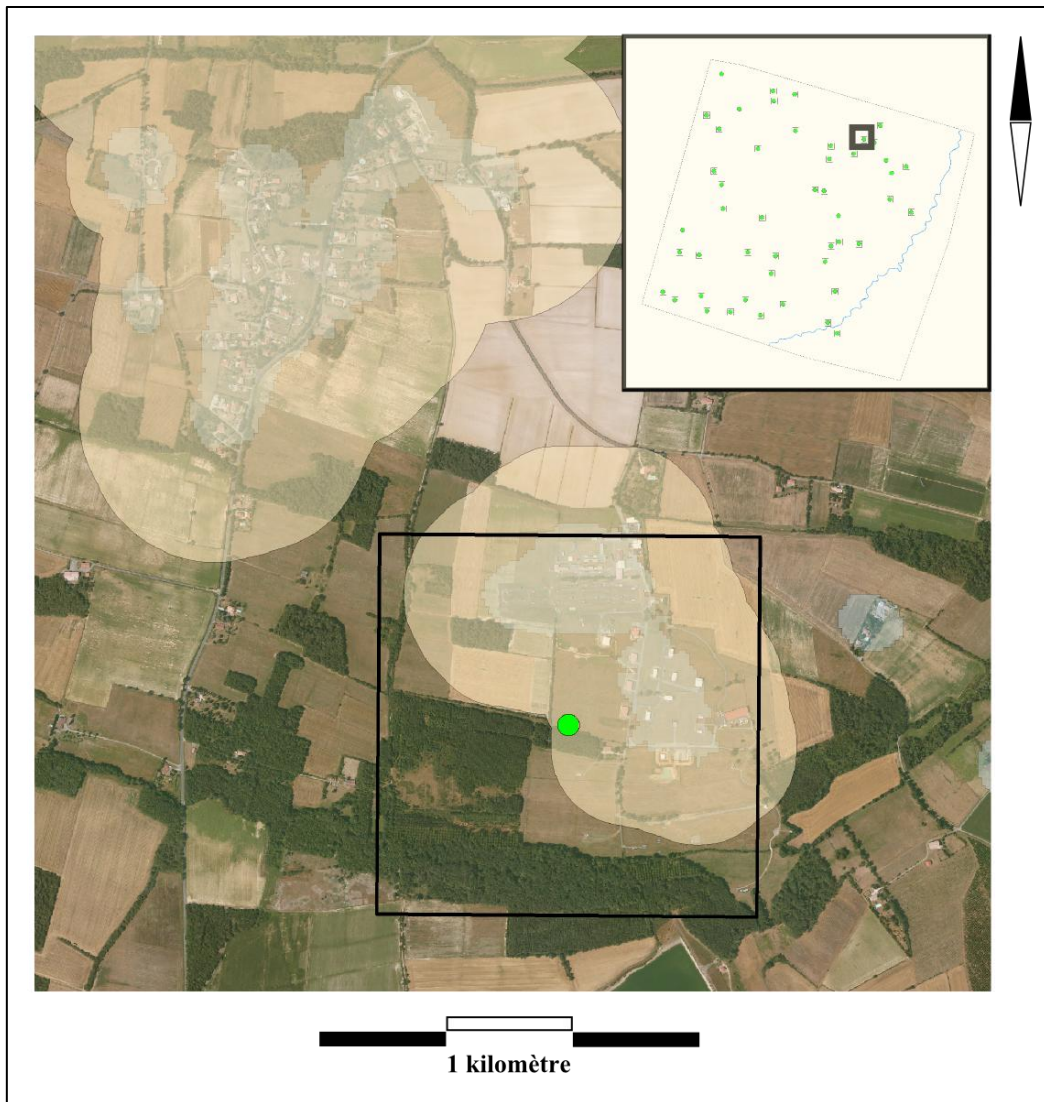


Figure 25 : Détail de la représentation cartographique visant à délimiter les zones d'agglomération urbaine. Les zones beige clair délimitent les zones d'habitation et les emprises en beige foncé délimitent les zones d'agglomération. Le point vert entouré d'un cadre représente le barycentre de l'un des 48 quadrats échantillonnés initialement (position indiquée dans l'encart situant l'aire d'étude). Le fond de carte basé sur la BD ORTHO de l'IGN permet de s'assurer de la pertinence du paramétrage utilisé.

Nous avons alors recalculé nos différents points d'échantillonnage en les situant au plus près à 700 mètres des zones d'agglomération (valeur de la demi-diagonale d'un carré de 1 kilomètre de côté). De cette manière, nos différents quadrats ne seront recouvrants avec aucune des zones d'agglomérations retenues. Cela présentera alors un double avantage: éviter de biaiser les analyses ultérieures par la présence d'un pourcentage élevé d'occupation du sol dédié au bâti et surtout redonner une certaine pertinence à nos rasters basés sur les indices d'hétérogénéité. Le fait d'avoir négligé le bâti a en effet conduit à biaiser le calcul de ces indices dans un rayon de 700 mètres autour de ces zones d'habitations. La réalisation, suite à la classification, d'une analyse majoritaire ainsi que l'élimination des parcelles de surface inférieure à 1 hectare a probablement limité les effets de ces erreurs de classification. Toutefois, une part non négligeable des points situés en agglomération reste mal classée. Cela apparaît facilement si l'on superpose la couche cartographique correspondant à l'occupation du sol par du contexte non agricole et la couche correspondant à ces zones d'agglomérations. Sur la base de cette procédure, nous sommes alors en mesure de proposer 12 points pour chacune des combinaisons des deux modalités d'indices d'hétérogénéité (figure 26). Un nouvel atlas répertoriant ces différents quadrats a été réalisé.

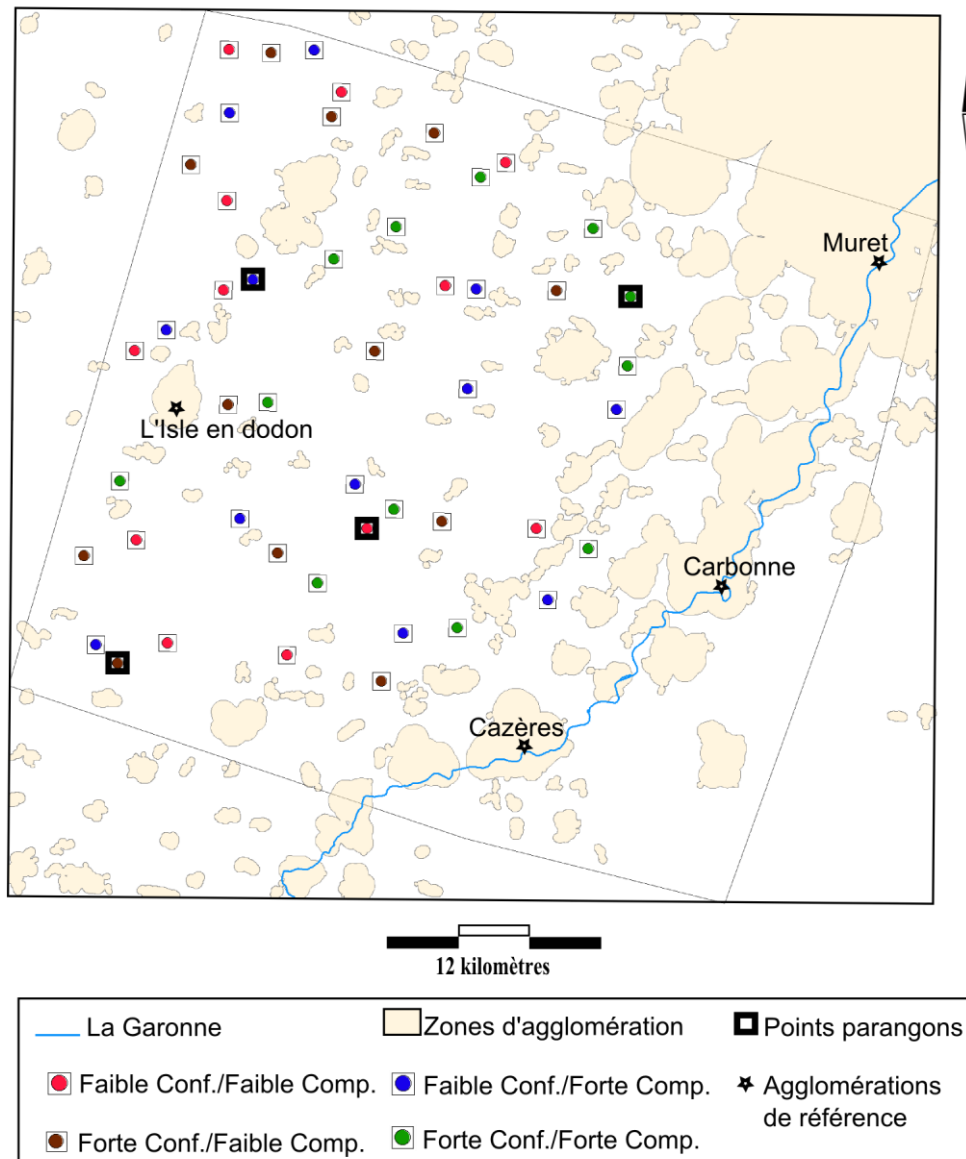


Figure 26 : Distribution spatiale des quadrats d'échantillonnage correspondant à chacune des combinaisons de modalités d'indices d'hétérogénéité de configuration et de composition (respectivement « Conf. » et « Comp. »). Les quatre quadrats dont le contour est le plus épais ont été choisi comme parangon de leur type.

La figure 27 présente la diversité des paysages des différentes combinaisons au travers de quatre parangons.



Figure 27 : Détail de l'occupation du sol des quatre parangons choisis pour illustrer les différentes combinaisons d'indice d'hétérogénéité sur le fond de photographie aérienne de 2006 (BD ORTHO, IGN). De la gauche vers la droite en haut puis en bas : Faible hétérogénéité de configuration-Faible hétérogénéité de composition, Forte hétérogénéité de configuration-Faible hétérogénéité de composition, Faible hétérogénéité de configuration-Forte hétérogénéité de composition et Forte hétérogénéité de configuration-Forte hétérogénéité de composition.

X. Conclusions générales et perspectives.

Notre objectif était de déployer, sur notre terrain d'étude des Coteaux de Gascogne, une méthode cartographique en s'inspirant des travaux récents d'une équipe de recherche Canadienne combinant des outils de géomatique, de télédétection spatiale et de géostatistique (Pasher *et al.*, 2011). A l'issue du stage, nous proposons 48 quadrats qui correspondent à 12 points d'échantillonnage par combinaison de modalités d'indices d'hétérogénéité et répondant au cahier des charges initial. Les différentes étapes de la chaîne de traitement sont maîtrisées et largement documentées, même si certaines améliorations restent envisageables voire

indispensables. L'apport de la géostatistique a été un point majeur pour la réussite de ce stage qui demandait de tester chemin faisant, à chaque étape, la qualité des résultats intermédiaires obtenus.

Avec le lancement officiel du projet FARMLAND en octobre 2012 et la mise en place des collaborations avec les chercheurs des différents sites d'études choisis, il faudra réappliquer toute la chaîne de traitement pour le site des Coteaux de Gascogne avec des images satellite plus récentes (images SPOT 5 2011 par exemple). Il faudra également améliorer la précision des classifications de l'étape 1 à partir desquelles découlent les résultats des étapes suivantes. Une étape préliminaire pourrait concerner la délimitation des principales zones d'agglomération urbaine et d'en réaliser un masque pour le métafichier d'images satellitaires. Il pourrait également être intéressant de réaliser la construction de zones d'apprentissage plus pures à partir des données RPG en utilisant des zones tampons d'érosion.

L'estimation de la distance minimale entre deux quadrats pourrait être affinée. Nous pouvons proposer un ajustement du calcul de la distance minimale entre deux centres de quadrats en prenant en compte la mesure de la diagonale des quadrats et non une distance légèrement supérieure à la longueur d'un quadrat tel que proposé dans le travail de Pasher *et al.*, 2011). Les conséquences dans le cadre des études d'écologie du paysage ont été mises en évidence par Eigenbord *et al.* (2011).

Il nous est apparu comme essentiel de ne pas attendre la fin de la chaîne de traitement pour effectuer une vérification terrain combinée au recours plus régulier aux orthophotographies couleur les plus récentes disponibles sur la zone d'étude. Ces vérifications devraient être systématiquement effectuées à la fin de chacune des grandes étapes de cette chaîne de traitement.

Nous proposons donc une nouvelle version de la chaîne de travail (figure 28) qui prend en compte les différentes recommandations évoqué ci-dessus.

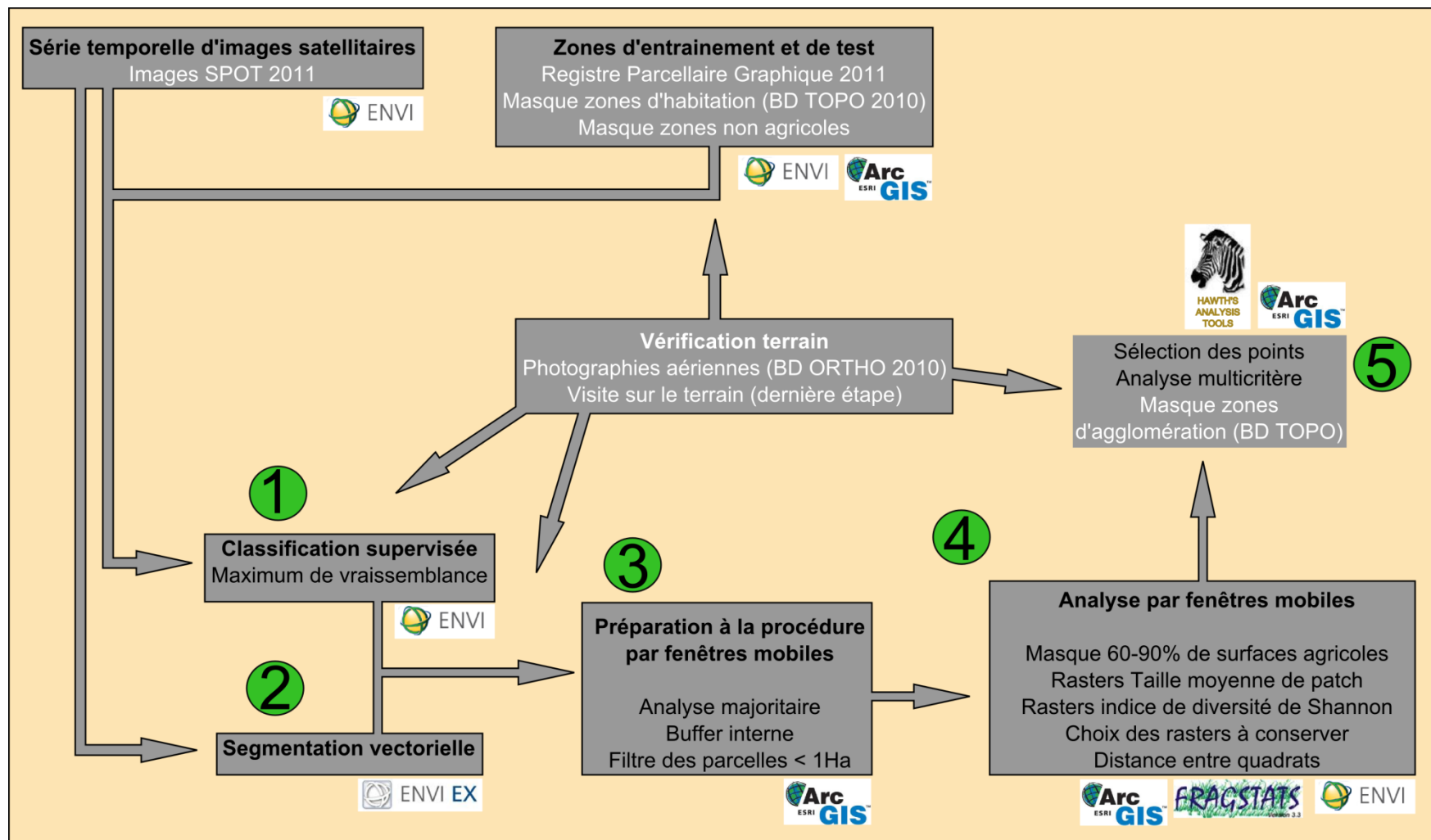


Figure 28: Nouvelle version de la chaîne de traitement. Les modifications suggérées apparaissent en blanc.

Ce stage aura également été l'occasion de développer un intérêt pour les analyses par fenêtres mobiles du logiciel FRAGSTATS v3.3 pour différentes études du laboratoire. Cela s'est traduit à brève échéance par la rédaction d'une fiche qualité sur l'utilisation des procédures par fenêtres mobiles avec ce logiciel (en annexe). A plus longue échéance, ce type d'analyse sera utilisé non seulement dans le projet FARMLAND, mais aussi pour traiter d'autres données dans le cadre d'autres projets de recherche nationaux (réseau INRA-SEBIOPAG) ou européens cherchant à relier l'étude du paysage et impact sur la biodiversité.

Une réunion de travail impliquant les différents partenaires européens du projet FARMLAND est déjà en cours de programmation pour l'automne 2011 afin de présenter les différentes étapes de sélection des sites et d'ores et déjà je suis invité à visiter le laboratoire de recherche de Lenore Fahrig à la Carlton University d'Ottawa pour échanger sur mon travail de stage et mon implication future dans le projet FARMLAND avec les différents partenaires européens. Une communication orale impliquant également Gérard Balent, Sylvie Ladet et Lenore Fahrig sera proposée pour les prochaines journées Françaises de l'écologie du paysage à Aix en Provence (IALE 2011 : <https://ialefrance2011.cemagref.fr/>).

A plan personnel, ce stage m'a permis de prendre pleinement possession de ma nouvelle activité dans le cadre de ma reconversion professionnelle tout en bénéficiant de l'expérience précédemment acquise. Le stage aura été l'occasion de transposer mon expérience de l'analyse de données acquise au fil de mes activités de Chargé de Recherche dans le monde de la géomatique. Cette dimension de l'analyse de données impliquant un recours aux analyses statistiques s'est révélée particulièrement enrichissante. L'usage de différents logiciels de géomatique, de télédétection, d'analyses statistiques dans une logique d'interopérabilité s'est également révélé riche d'enseignement tout en y mêlant le plaisir de l'activité de recherche issu de la confrontation de perspectives différentes sur un sujet donné. Ce stage m'aura permis de m'initier au travail de l'ingénieur en me mettant au service des chercheurs pour répondre à des problèmes ponctuels temporellement parlant. Cette grande réactivité demandée à l'ingénieur est indéniablement un des aspects de ce travail qui m'a le plus séduit.

Ce nouveau départ me conduit à rester au moins pour un quadriennal dans l'UMR DYNAFOR, soit la durée du programme FARMLAND que je pourrai suivre en qualité de correspondant technique en plus des autres travaux auxquels je serai amené à collaborer dans cette unité. La qualité des conditions de travail ainsi que les excellents rapports entre collègues dans cette unité m'offrent donc un contexte de travail particulièrement épanouissant que je souhaite pérenniser dans le cadre des fonctions d'ingénieur.

XI. Références bibliographiques.

- Altieri, M.A.** (1999). The ecological role of biodiversity in agroecosystems. *Agric. Ecosyst. Environ.*, 74, 19–31.
- Arias M. et Barry N.** (2011) Cartographie de l'occupation du sol d'un paysage agricole à partir de données satellitaires et agricoles. Rapport de miniprojet SIGMA. 44pp.
- Benton T.G., Vickery J.A. et Wilson, J.D.** (2003). Farmland biodiversity: is habitat heterogeneity the key? *Trends Ecol. Evol.*, 18, 182–188.



- Burton R.J.F., Kuczera C. et Schwarz, G.** (2008). Exploring farmers cultural resistance to voluntary agri-environmental schemes. *Sociol. Ruralis*, 48, 16–37.
- Davidson A.D., Hamilton M.J., Boyer A.G., Brown J.H., et Ceballos G.** (2009). Multiple ecological pathways to extinction in mammals. *Proc. Nat. Acad. Sci.* 106, 1702-1705.
- Duelli P.** (1992). Mosaikkonzept und Inseltheorie in der Kultur-landschaft. *Verh. Ges. für Ökologie* 21, 379-384.
- Duelli P.** (1997). Biodiversity evaluation in agricultural landscapes: an approach at two different scales. *Agric. Ecosyst. Environ.*, 62,81–91.
- Duelli P. et Obrist, M.K.** (2003). Regional biodiversity in an agricultural landscape: the contribution of seminatural habitat islands. *Basic Appl. Ecol.*, 4, 129–138.
- Dunning J.B., Danielson B.J. et Pulliam H.R.** (1992). Ecological processes that affect populations in complex landscapes. *Oikos*, 65, 169–175.
- Eigenbrod F., Hecnar S.J., et Fahrig L.** (2011) Sub-optimal study design has major impacts on landscape-scale inference. *Biol. Cons.*, 144, 298-305.
- Fahrig L.** (2003). Effects of habitat fragmentation on biodiversity. *Annu. Rev. Ecol. Syst.*, 34, 487–515.
- Fahrig L. et Nuttle W.K.** (2005). Population ecology in spatially heterogeneous environments. In: *Ecosystem Function in Heterogeneous Landscapes* (eds Lovett, G.M., Jones, C.G., Turner, M.G. & Weathers, K.C.). Springer-Verlag, New York, pp. 95–118.
- Fahrig L., Baudry J., Brotons L., Burel F., Crist T.O., Fuller R.J., Sirami C., Siriwardena G.M. et Martin J.-L.** (2011) Functional heterogeneity and biodiversity in agricultural landscapes. *Ecol. Lett.*, 14, 101-112.
- Foley J.A., DeFries R., Asner G.P., Barford C., Bonan G., Carpenter S.R., Chapin F.S., Coe M.T., Daily G.C., Gibbs H.K., Helkowski J.H., Holway T., Howard E.A., Kucharik C.J., Monfreda C., Patz J.A., Prentice I.C., Ramankutty N. et Snyder P.K.** (2005). Global consequences of land use. *Science*, 309, 570–574.
- Geiger F., Bengtsson J., Berendse F., Weisser W. W., Emmerson M., Morales M. B., Ceryngier P., Liira J., Tsharntke T., Winqvist C., Eggers S., Bommarco R., Prt, T., Bretagnolle V., Plantegenest M., Clement L. W., Dennis C., Palmer C., Oñate J. J., Guerrero I., Hawro V., Aavik T., Thies C., Flohre A., Hnke S., Fischer C., Goedhart P. W. et Inchausti P.** (2010). Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland *Basic Appl. Ecol.*, 11, 97-105.
- Gurr G.M., Wratten S.D. et Luna J.M.** (2003). Multifunction agricultural biodiversity: pest management and other benefits. *Basic. Appl. Ecol.* 4, 107-116.
- Le Roux X., Barbault R., Baudry J., Burel F., Doussan I., Garnier E., Herzog F., Lavorel S., Lifran R., Roger-Estrade J., Sarthou J.P., et Trommetter M. (éditeurs).** (2008). *Agriculture et biodiversité. Valoriser les synergies.* Expertise scientifique collective, rapport, INRA (France).
- McGarigal K., Cushman S.A., Neel M.C., Ene E.** (2002). FRAGSTATS: Spatial pattern analysis program for categorical maps. Comp. software prog. Univ. Mass., Amherst. www.umass.edu/landeco/research/fragstats/fragstats.html
- Odum E.P.** (1971) *Fundamentals of Ecology*. Sanders College Publications, Philadelphia, 3rd ed.
- Ouin A. et Burel F.** (2002). Influence of herbaceous elements on butterfly diversity in hedgerow agricultural landscape. *Agric. Ecosyst. Environ.*, 93, 45–53
- Pasher J, Fahrig L, King D, Lindsay K, et Mitchell S.** In preparation. Selecting sample landscapes for a landscape-scale mensurative experiment.
- Pfiffner L. et Luka H.** (2003). Effects of low-input farming systems on carabids and epigeal spiders – a paired farm approach. *Basic Appl. Ecol.*, 4, 117–127.

- Polis G.A., Power M.E. et Huxel G.R. (eds)** (2004). *Food Webs at the Landscape Level*. University of Chicago Press, Chicago.
- Smart S.M., Thompson K., Marrs R.H., Le Duc M.G., Maskell L.C. et Firbank L.G.** (2006). Biotic homogenization and changes in species diversity across human-modified ecosystems. *Proc. R. Soc. B* 2006 273, 2659-2665.
- Smith P., Martino D., Cai Z., Gwary D., Janzen H., Kumar P., McCarl B., Ogle S., O'Mara F., Rice C., Scholes B., Sirotenko O., Howden M., McAllister T., Pan G., Romanenkov V., Schneider U., Towprayoon S., Wattenbach M. et Smith J.** (2008). Greenhouse gas mitigation in agriculture. *Phil. Trans. R. Soc. B* 2008 363, 789-813
- Srinivasan U.T., Carey S.P., Hallstein E., Higgins P.A.T., Kerr A.C., Koteen L.E., Smith A.B., Watson R., Harte J., and Norgaard R.B.** (2008). The debt of nations and the distribution of ecological impacts from human activities. *Proc. Nat. Acad. Sci.* 105, 1768-1773.
- Thies C. et Tschardtke T.** (1999). Landscape structure and biological control in agroecosystems. *Science*, 285, 893–895.
- Tschardtke T., Klein A.M., Kruess A., Steffan-Dewenter I. et Thies C.** (2005). Landscape perspectives on agricultural intensification and biodiversity: ecosystem service management. *Ecol. Lett.*, 8, 857–874.
- Zalasiewicz J., Williams M., Haywood A. and Ellis M.** (2011). The Anthropocene: a new epoch of geological time. *Phil. Trans. R. Soc. A* 369, 835-841.
- Beyer H.L.** 2004. Hawth's Analysis Tools for ArcGIS. Available at <http://www.spatial ecology.com/htools>.

ANNEXES

ANNEXE 1

FICHE QUALITE UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	

1. Objet et domaine d'application:

Illustrer les différentes étapes d'une procédure de fenêtres mobiles avec le logiciel FRAGSTATS.

2. Documents de référence :

Guide de l'utilisateur de FRAGSTATS :

<http://www.umass.edu/landeco/research/fragstats/documents/User%20guidelines/User%20guidelines%20content.htm>

3. Hygiène et sécurité :

RAS

4. Autres matériels ou documents nécessaires :



Liste des documents ou informations ou matériels
:

Disponibles auprès de :

◆ Logiciel SIG raster (ArcGis 9.3.1 par ex.)

◆ Sylvie Ladet

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	

- ◆ Logiciel FRAGSTATS 3.3 (téléchargeable gratuitement)
- ◆ Editeur de texte type Bloc-note ou Wordpad
- ◆ Sylvie Ladet
- ◆

5. Conditions et/ou contraintes à respecter :

Conduite à tenir face aux problèmes principaux :



Problèmes ou contraintes	Conduite à tenir :
◆	◆
◆	◆
◆	◆

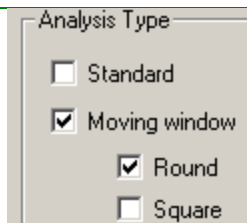
6. Contenu de l'instruction

La procédure de fenêtre mobile se comporte un peu comme un filtre de lissage. Une fenêtre (de forme ronde ou carré dont la taille est à déterminer) parcourt le paysage et va donner une valeur au pixel central de la fenêtre qui dépendra de la métrique choisie (valeur moyenne, médiane, indice de diversité...).

Le logiciel Fragstats propose un nombre considérable de métriques et à des niveaux scalaires variables (classe, patch, paysage). Les métriques sont aussi appelées indicateurs métriques de paysage c'est-à-dire indices qui sont calculés à partir d'une carte des éléments du paysage (le plus souvent une carte d'occupation du sol) et utilisés pour évaluer les différentes propriétés ou fonctions des paysages.

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	




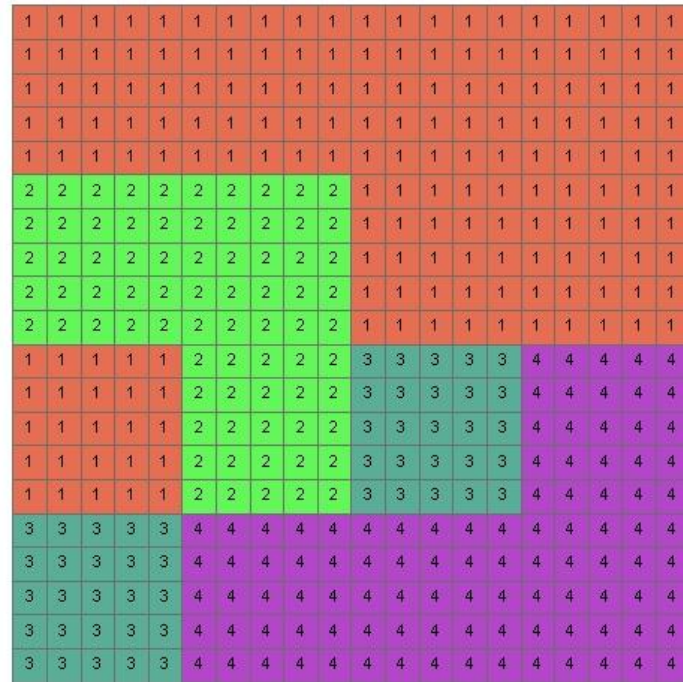
Pour l'exemple présenté ici nous utiliserons les analyses de type par fenêtre mobile uniquement l'indice de diversité de Shannon, métrique qui est disponible à l'échelle du paysage (i.e. la fenêtre mobile).

6.1 Les données de départ :

Pour le cas qui nous concerne, un raster (représentant un paysage virtuel très simplifié) a été créé sur Arcgis grâce à l'outil Hawth's Tools et un compte numérique a été attribué aux pixels afin d'avoir un environnement peu complexe et didactique. Ce raster carré a une dimension de 2000 mètres de côté et une résolution à 100 mètres. Il comporte donc 400 cellules. 4 modalités d'occupation du sol ont été choisies : 1, 2, 3, 4. Toutefois la dernière modalité sera recodée par la suite en *Background* avec un compte numérique de -99.

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			



 UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	



6.2 Le cas du *background*, des limites et des bordures dans Fragstat :

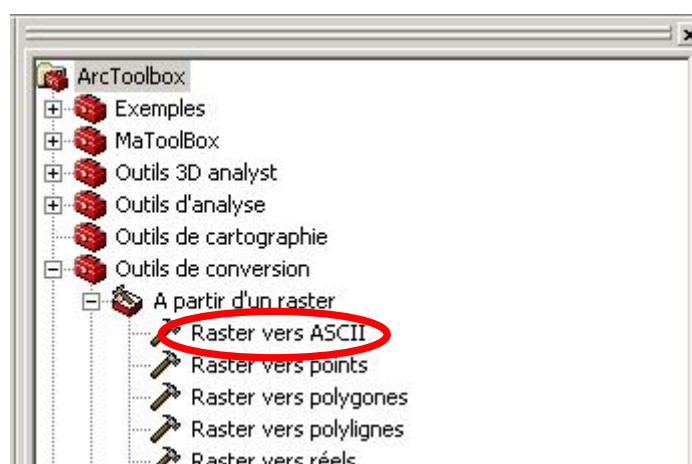
Dans Fragstats, les pixels vont avoir un compte numérique dont la valeur pourra être un nombre entier ou réel. Toutefois, certains pixels pourront ne pas avoir de valeur (« trou » dans l'image). Ces pixels recevront une valeur particulière (souvent 99) comme compte numérique pour identifier ces données manquantes. Il sera alors possible soit d'intégrer ces pixels dans les calculs ultérieurs (compte numérique positif) soit de les ignorer (compte numérique négatif). Ce procédé permettra également d'éliminer d'éventuels effets de bord en créant une bordure à l'image avec un compte numérique de -99. Dans notre exemple, nous n'utiliserons pas ce dernier procédé et le *background* sera retiré des calculs. Les pixels dont le compte numérique a une valeur de 4 seront ainsi recodés avec la valeur -99.

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	

6.3 Format d'import du raster dans Fragstats :

Fragstats accepte différents formats de données, mais le plus simple est d'utiliser le format ASCII. Il convient donc de convertir le raster de départ avec l'outil de conversion « raster vers ASCII » de l'ArcToolbox d'Arcgis.



Une fois cette opération effectuée, le raster devient le fichier texte suivant :

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			



UMR 1201 DYNAFOR

Instruction SIG

UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS

Réf. : **INS-SIG-05**
Version : 01
date : 10/05/2011



```

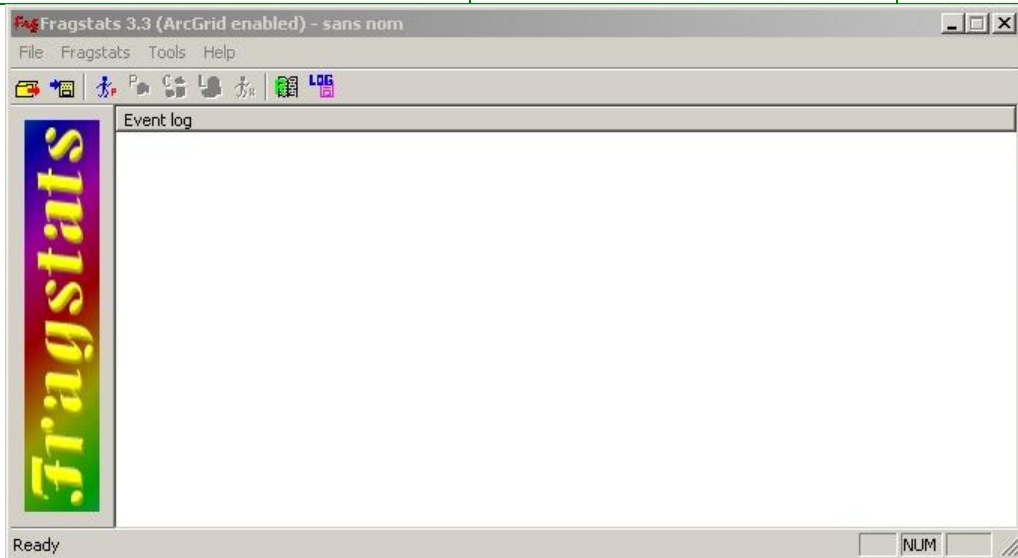
Sans titre - Bloc-notes
Fichier Edition Format Affichage ?
ncols          20
nrows          20
xllcorner      468000
yllcorner      1842000
cellsize       100
NODATA_value   -99
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 -99 -99 -99 -99 -99
1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 -99 -99 -99 -99 -99
1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 -99 -99 -99 -99 -99
1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 -99 -99 -99 -99 -99
1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 -99 -99 -99 -99 -99
3 3 3 3 3 3 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99
3 3 3 3 3 3 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99
3 3 3 3 3 3 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99
3 3 3 3 3 3 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99 -99

```

Pour l'importer dans Fragstats, il faut retirer l'en-tête du fichier texte obtenu (soit les 6 premières lignes) et l'on obtient le fichier texte sur lequel Fragstats va travailler :

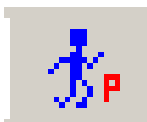
	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	





6.5 Création d'un fichier de paramètres généraux:

La sélection de la troisième icône en partant de la gauche permet de lancer l'interface de paramétrage suivante :



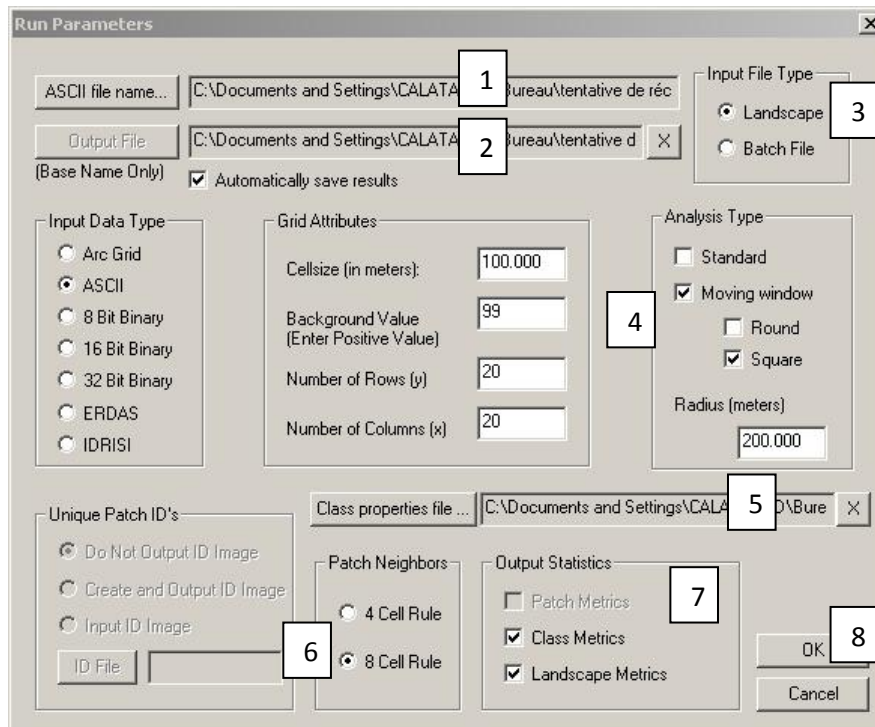
Etapas à suivre :

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	

1°) choisir le fichier sur lequel nous allons travailler.



2°) donner un nom au fichier de sortie. Dans notre cas, le paramétrage suivant a été choisi :



3°) Nous travaillons sur un seul fichier, mais le logiciel offre la possibilité de travailler avec une procédure de *Batch* (travail en série sur plusieurs fichiers).

4°) Nous utilisons un fichier ASCII dont la taille de cellule est de 100 mètres, la valeur absolue du compte numérique codant le *Background* est 99 et le fichier comporte 20 lignes et 20 colonnes. Nous utiliserons une fenêtre mobile de forme carrée de 500 mètres de côté. Pour paramétrer cette taille de fenêtre, Fragstats met en œuvre un paramètre appelé radius qui correspond à une distance reportée de part et d'autre de la cellule centrale. La taille de la


	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	<p align="center">Instruction SIG</p> <p align="center">UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS</p>	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
---	---	---

fenêtre sera donc égale à la somme de la taille de la cellule centrale (la résolution du pixel : 100 mètres) et de deux fois la valeur du radius. Dans notre cas : 100+200+200=500 mètres.

5°) Pour l'analyse, Fragstats a besoin d'un fichier texte permettant de décrire les différentes modalités (classes) composant le fichier ASCII. Il s'agit d'un fichier texte utilisant la virgule comme séparateur qui va lister l'identificateur de classe (le compte numérique attribué), le nom de la classe, son statut (participant ou non à l'analyse : respectivement valeur t (pour « true ») ou f (pour « false »)) et si la classe correspond ou non à du *background* (respectivement valeur t (pour « true ») ou f (pour « false »)).

Ce fichier prend dans notre cas la forme suivante :



```

class_prop - Bloc-notes
Fichier Edition Format Affichage ?
-99, background, f, t
1, nonagri, t, f
2, cereales, t, f
3, colza, t, f



```

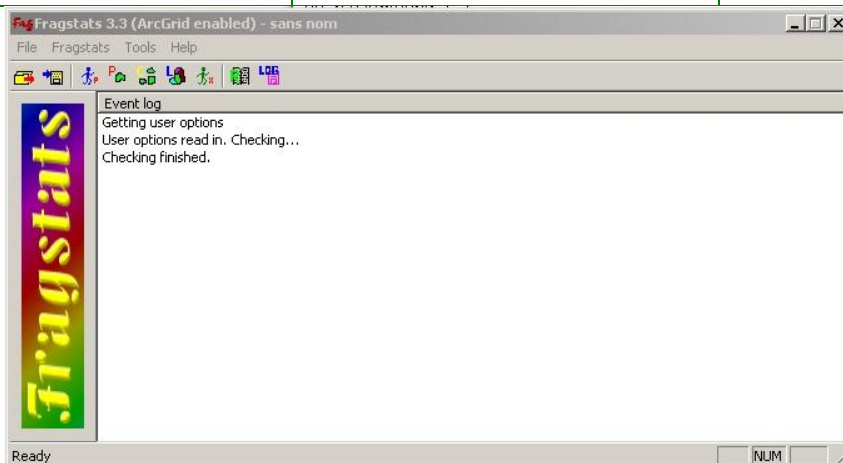
6°) Il convient ensuite de choisir la règle de voisinage des cellules (utilisant ou non la diagonale pour le définir). Dans le paramétrage à 4 cellules, seules les pixels en contact par un côté sont utilisés (forme de +) alors que dans le cas du paramétrage à 8 cellules, tous les pixels sont utilisés.

7°) Il ne reste plus qu'à choisir le type de métriques que l'on souhaite obtenir. Dans notre cas, nous voulons utiliser l'indice de Shannon qui est disponible à l'échelle du paysage (Landscape metrics).

8°) Une fois tous ces paramètres réglés, il est possible de valider, ce qui permet de rebasculer automatique sur la fenêtre de démarrage avec de nouvelles options.

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	




Des nouvelles icônes (quatrième jusqu'à la septième de la gauche vers la droite), permettent respectivement de paramétrer les métriques au niveau du *patch*, de la *classe* ou du *landscape* (paysage) selon ce qui a été choisi dans la fenêtre de paramétrage précédente (normalement dans notre cas, seule la fenêtre « paysage », sixième en partant de la gauche, devrait être activée) et enfin de faire exécuter le calcul demandé.

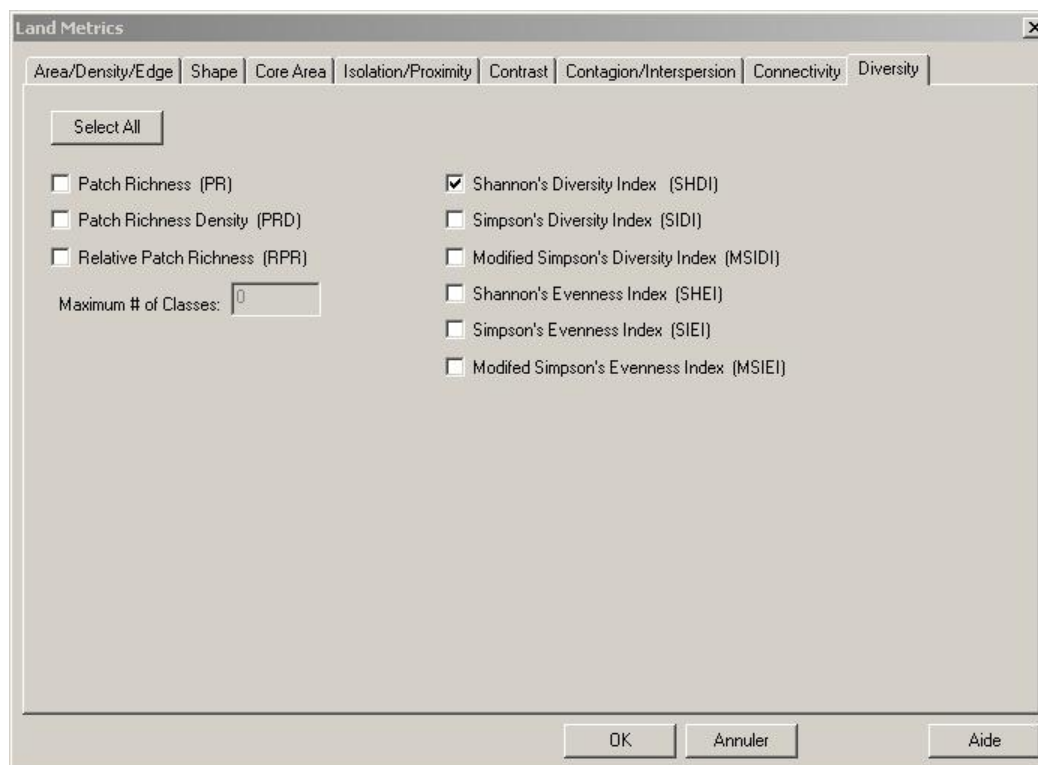
6.6 Création d'un fichier de paramètres de métriques:

La sélection par clic sur l'icône de paramétrage des métriques de paysage ouvre une nouvelle interface :




	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

 UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	





Nous souhaitons utiliser l'indice de diversité de Shannon, celui-ci se trouve dans le dernier onglet « *Diversity* ».

Une fois le champ correspondant coché, il est possible de revenir à l'interface de départ et de faire exécuter le calcul demandé en cliquant sur la septième icône en partant de la

gauche. 

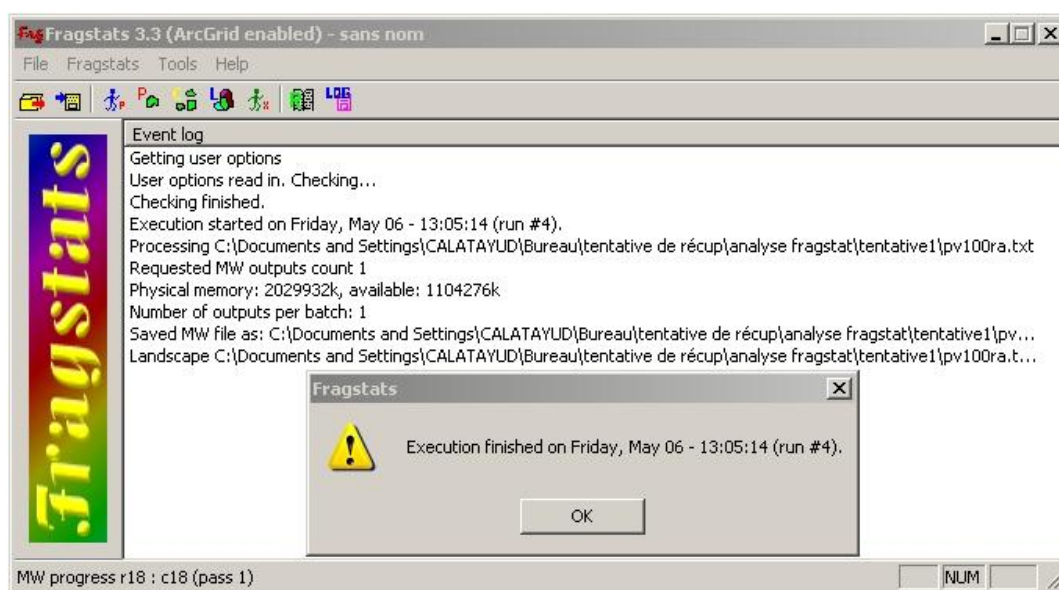
	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	

Le calcul peut demander du temps (voire beaucoup de temps...) selon la complexité du raster de départ et le nombre de métriques demandées. Notre exemple correspond toutefois à une situation très simple, et l'exécution du calcul est rapide.



6.7 Mise en forme des fichiers de sortie :

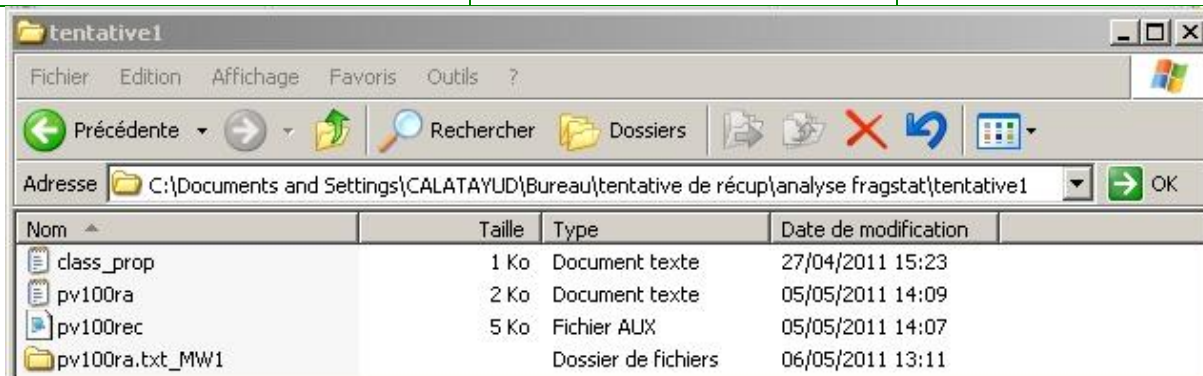
Une fois ce calcul terminé, un message le signale :



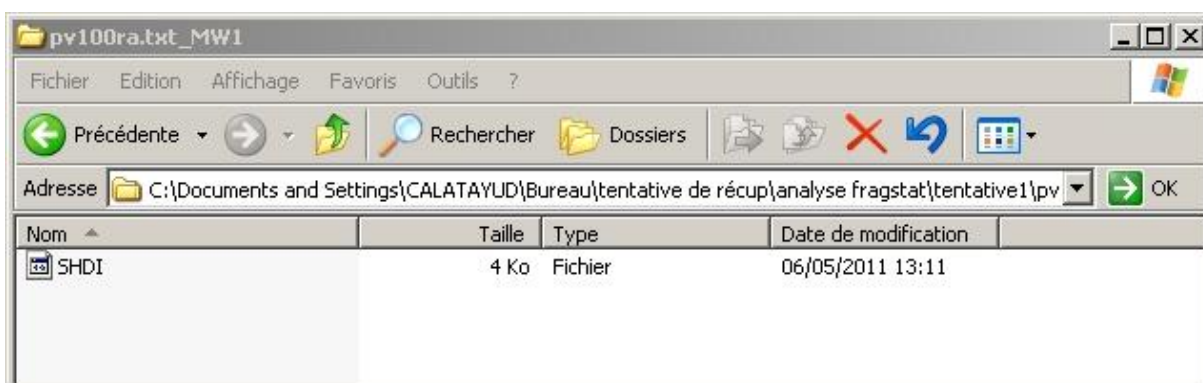
Un nouveau dossier a été créé dans notre répertoire de travail. Il porte le nom du fichier ascii de départ suivi de « _MW* » (pour *moving window*) où le caractère * est remplacé par un chiffre. Il est ainsi possible de faire plusieurs exécutions de suite en changeant quelques paramètres et de classer les différentes opérations dans leur ordre chronologique. Dans notre cas, ce dossier s'appelle pv100ra.txt_MW1.

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	





Ce dossier contient les fichiers de sortie sous un format lisible dans un éditeur de texte ou dans Excel.



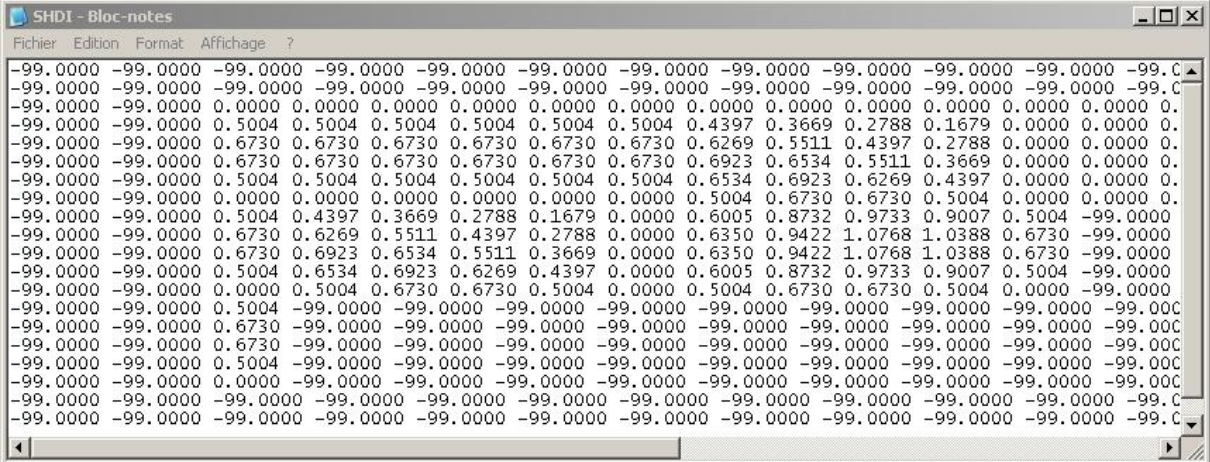
Dans notre cas il ne contient qu'un seul fichier appelé SHDI (pour *Shannon Diversity Index*).

Ce fichier va contenir les données permettant de réaliser une carte d'analyse de voisinage utilisant l'indice de diversité de Shannon pour une fenêtre mobile de 500 mètres de côté. Très peu de travail reste à faire pour obtenir une carte raster avec ces données sous Arcgis!

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	

Comme cela a été dit précédemment, le fichier de sortie peut être lu par un éditeur de texte.



```


-99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
-99.0000 -99.0000 0.5004 0.5004 0.5004 0.5004 0.5004 0.5004 0.5004 0.5004 0.4397 0.3669 0.2788 0.1679 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
-99.0000 -99.0000 0.6730 0.6730 0.6730 0.6730 0.6730 0.6730 0.6730 0.6730 0.6269 0.5511 0.4397 0.2788 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
-99.0000 -99.0000 0.6730 0.6730 0.6730 0.6730 0.6730 0.6730 0.6730 0.6730 0.6923 0.6534 0.5511 0.3669 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
-99.0000 -99.0000 0.5004 0.5004 0.5004 0.5004 0.5004 0.5004 0.5004 0.5004 0.6534 0.6923 0.6269 0.4397 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
-99.0000 -99.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.5004 0.6730 0.6730 0.5004 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
-99.0000 -99.0000 0.5004 0.4397 0.3669 0.2788 0.1679 0.0000 0.6005 0.8732 0.9733 0.9007 0.5004 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.6730 0.6269 0.5511 0.4397 0.2788 0.0000 0.6350 0.9422 1.0768 1.0388 0.6730 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.6730 0.6923 0.6534 0.5511 0.3669 0.0000 0.6350 0.9422 1.0768 1.0388 0.6730 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.5004 0.6534 0.6923 0.6269 0.4397 0.0000 0.6005 0.8732 0.9733 0.9007 0.5004 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.0000 0.0000 0.5004 0.6730 0.6730 0.5004 0.0000 0.5004 0.6730 0.6730 0.5004 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.5004 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.6730 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.5004 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 0.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000
-99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000 -99.0000

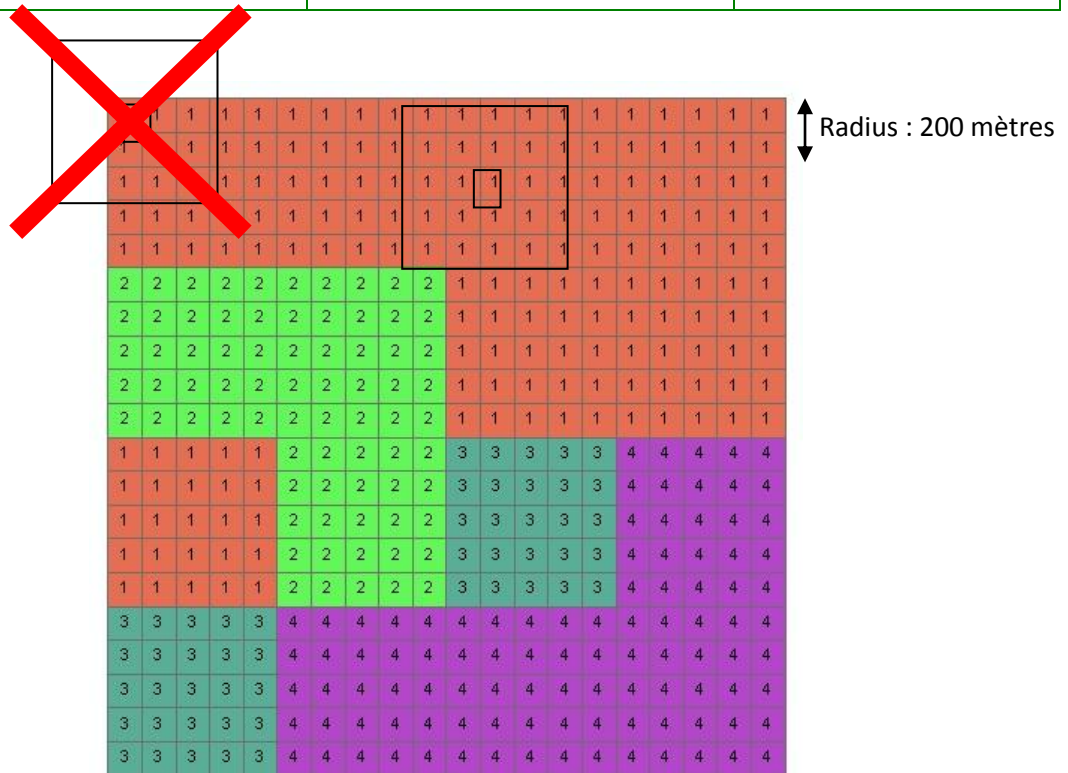
```

On reconnaît la structure d'une matrice de 20 lignes et 20 colonnes dans les cases de laquelle apparaît le compte numérique des pixels avec la valeur de l'indice de Shannon.

A parte : On notera que la taille de la matrice impose un effet de bord. Seuls les pixels pouvant prendre la place du pixel central de la fenêtre mobile peuvent participer au calcul et les autres prennent une valeur de *Background* (-99). Sur la figure suivante, un carré transparent à bord noir illustre la taille d'une fenêtre et un petit carré sa cellule centrale. On s'aperçoit que les pixels pris en compte sont ceux situés au minimum à 2 pixels du bord, soit une distance de 200 mètres, la valeur du radius dans notre cas. Dans le cas des pixels situés trop près du bord, la fenêtre mobile ne contient pas suffisamment de pixels pour exister.

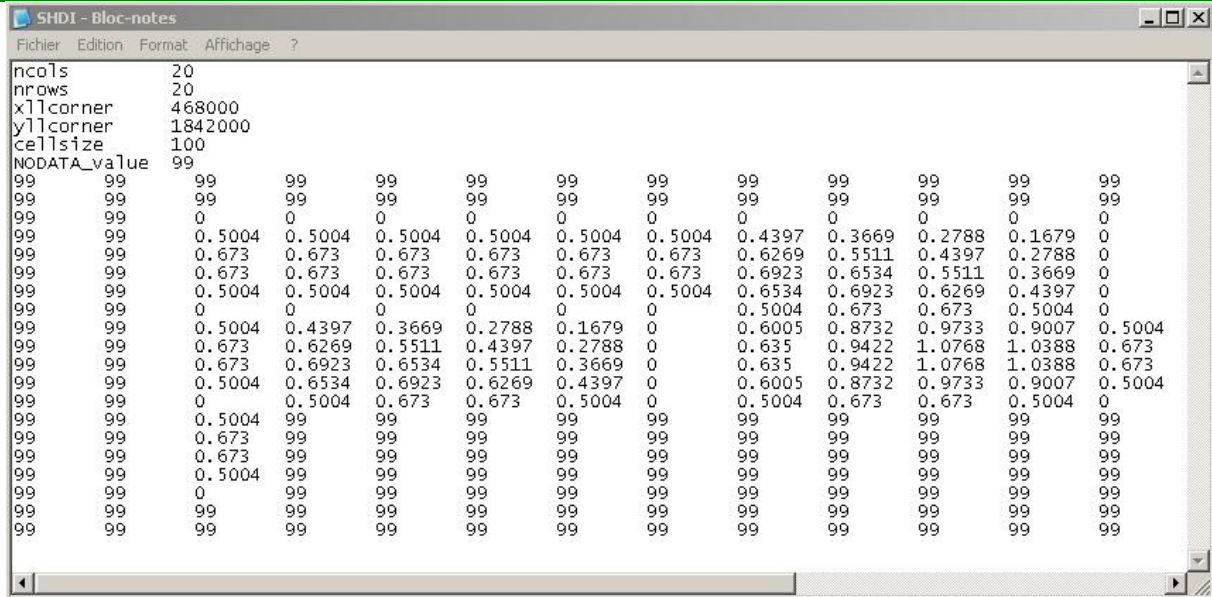
	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

 UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	



La matrice obtenue en sortie ayant la même structure que celle du fichier ASCII de départ, on peut lui attribuer le même en-tête (en spécifiant que les données manquantes sont codées par la valeur 99) et l'importer dans ArcGis.

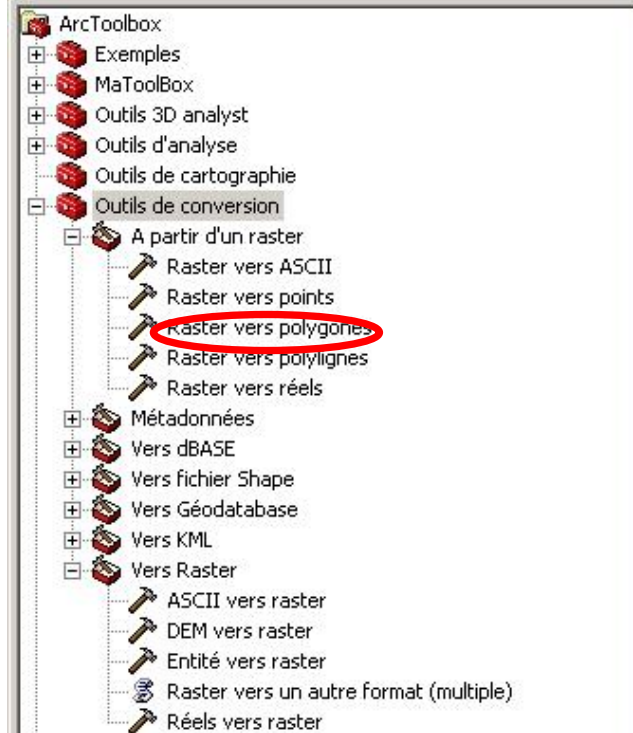
	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			



6.8 Import dans ArcGis :



L'import dans Arcgis se fait grâce à l'outil de conversion « ASCII vers Raster » de l'ArcToolbox :

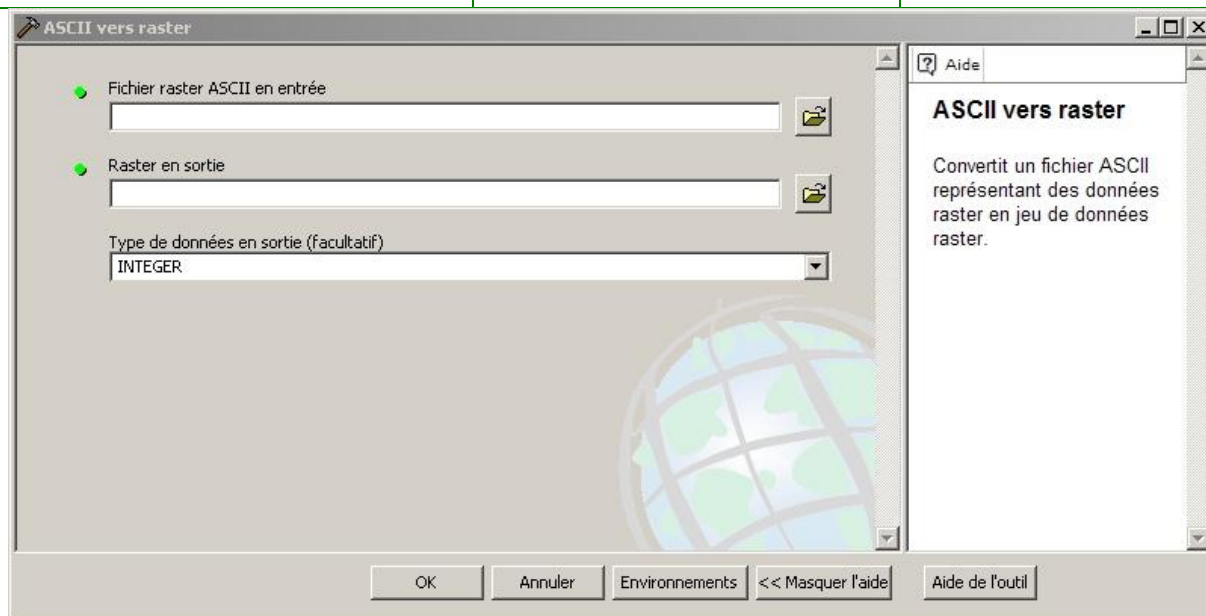
	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			



La sélection de cet outil ouvre une nouvelle fenêtre dans laquelle il suffit de spécifier le nom du fichier ASCII en entrée et celui du raster en sortie en précisant toutefois le type du compte numérique : nombre entier ou réel (*Integer* ou *Float*). Dans notre cas, c'est évidemment un compte numérique en virgule flottante qu'il convient de choisir.


	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

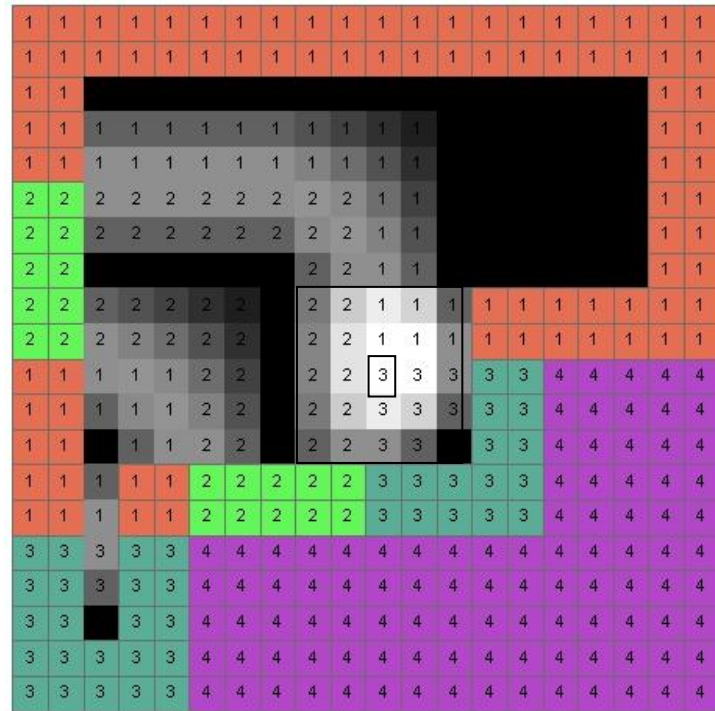
  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	



Une fois le calcul terminé, nous disposons d'une nouvelle couche raster superposable à notre raster de départ:

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			



 UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	



Valeur de l'Indice de Shannon
Elevée : 1.0768
Faible : 0

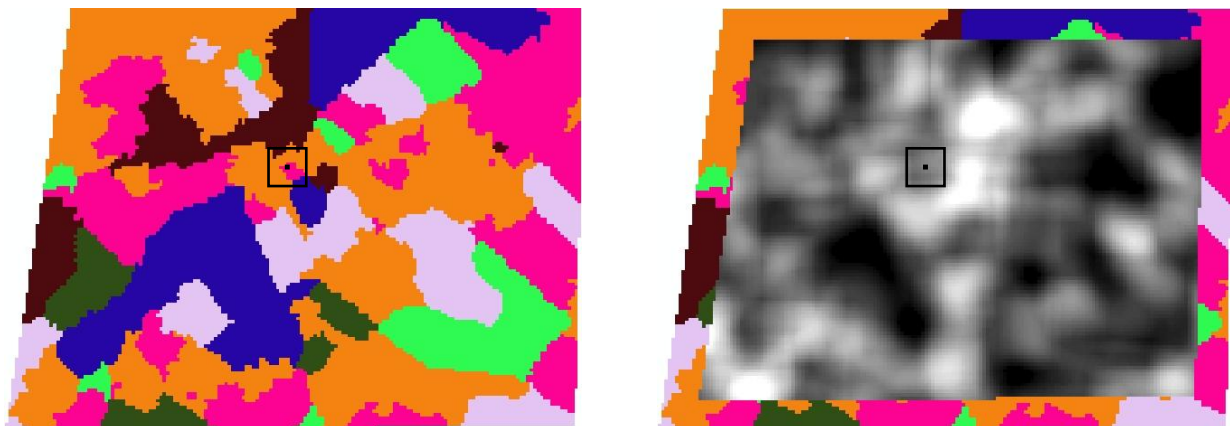
Ce nouveau raster décrit l'indice de diversité de Shannon calculé pour chaque pixel sur une surface de 25 hectares (représenté comme précédemment par un carré transparent à bord noir de 5 pixels de côté). Sur cette dernière figure, les valeurs fortes de l'indice de Shannon sont illustrées par une couleur claire (attention, les valeurs présentées illustrent le compte numérique du raster de départ et pas la valeur de l'indice de Shannon). Ces zones claires représentent les zones d'interface entre différents types d'occupation du sol.

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			



  UMR 1201 DYNAFOR	Instruction SIG	Réf. : INS-SIG-05 Version : 01 date : 10/05/2011
	UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS	

6.9 Exemple un peu plus réaliste :

La figure suivante montre un extrait d'une classification d'occupation du sol de 2009 à partir d'une image SPOT. Sept modalités d'occupation du sol ont été retenues par la classification (une couleur égale une modalité). Le raster de départ a une taille de 168 colonnes pour 123 lignes et présente une résolution à 10 mètres. Nous avons choisi une taille de fenêtre mobile carrée de 200 mètres de côté soit un radius de 95 mètres (un carré transparent à bord noir illustre la taille d'une fenêtre et un point sa cellule centrale). La carte de départ et la carte de l'indice de diversité se présentent sous la forme suivante :



	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

  UMR 1201 DYNAFOR	<p style="text-align: center;">Instruction SIG</p> <p style="text-align: center;">UTILISATION DES FENETRES MOBILES AVEC LE LOGICIEL FRAGSTATS</p>	<p>Réf. : INS-SIG-05 Version : 01 date : 10/05/2011</p>
---	---	--

7. Liste de diffusion et si nécessaire niveau de confidentialité :

Pour DYNAFOR et extérieur

8. Sauvegarde et archivage de cette instruction :

Une copie de ce document sera archivée sur le gestionnaire de documents « GDO » de l'UMR DYNAFOR accessible à partir de l'adresse suivante (Login et mot passe LDAP):

https://efpa.rennes.inra.fr/gdo_DYNAFOR.

	Rédacteur	Vérificateur	Approbateur
Nom :	François Calatayud	Sylvie Ladet	Laurent Burnel
Fonction :	Chargé de Recherche	Ingénieure d'étude	ANQ
Visa :			

ANNEXE 2

**PRESENTATION DES METRIQUES
INDICE DE DIVERSITE DE SHANNON
ET TAILLE MOYENNE DE TACHE**

Présentation des métriques

Indice de diversité de Shannon et Taille moyenne de tache

Contexte : Lors de notre travail, nous avons fait intervenir, à plusieurs reprises, différentes métriques calculées par le logiciel FRAGSTATS v3.3. Parmi celles-ci, Fahrig *et al.* (2010) proposent d'utiliser l'indice de diversité de Shannon comme un indice de l'hétérogénéité de composition et la taille moyenne de patch comme indice de l'hétérogénéité de configuration. Ces métriques sont alors calculées dans l'emprise d'une fenêtre mobile et le résultat est attribué à la cellule centrale sur la base du voisinage de celle-ci.

L'indice de diversité de Shannon est issu initialement de la théorie de l'information qui pose les bases mathématiques de la quantification de l'information contenue dans un message. L'indice de diversité de Shannon se présente comme la quantité moyenne d'information contenue dans un message basé sur différents éléments k dont le nombre peut varier de 1 à n . Cet indice se calcule suivant la formule :

$$H_s(p) = - \sum p_k \ln(p_k)$$

Dans cette formule, p_k représente la probabilité d'apparition (ou la fréquence d'occurrence) de chaque élément k_i . Ainsi, la valeur de cet indice dépendra non seulement du nombre d'éléments différents mais aussi de leur fréquence d'apparition. Il est également possible d'utiliser cet indice pour évaluer la diversité d'éléments présents sur une emprise donnée, ce qui a popularisé l'usage de cet indice en écologie. C'est dans ce cadre que FRAGSTATS v3.3 effectue ce calcul. Pour cela, ce calcul se base sur un raster comportant différentes modalités d'occupation du sol qui constitueront les différents éléments k_i . Le logiciel détermine le nombre de pixels de chaque modalité, leur fréquence et utilise ces valeurs pour calculer l'indice.

La taille moyenne de tache de végétation va être calculée à partir du même type de raster que l'indice de Shannon, et son calcul va être basé non seulement sur le nombre de pixels de chaque modalité, mais aussi sur leur taille. Le raster est défini par le compte numérique des pixels (la modalité d'occupation du sol), mais aussi par une résolution spécifiant que chacun d'entre eux aura la même taille (10 mètres dans notre cas). Autrement dit, chaque pixel correspondra à 100 mètres carrés d'occupation du sol. Il est dès lors possible de calculer la surface de chaque tache d'occupation du sol (nombre de pixels de chaque tache multiplié par 100 pour une surface en mètres carrés) puis la surface moyenne (somme des différentes surfaces calculées précédemment divisée par le nombre de taches). Cet indice sera alors directement relatif à la fragmentation du paysage en parcelles différentes.

ANNEXE 3

PRESENTATION DE L'INDICE D'AUTOCORRELATION DE MORAN

Présentation de l'indice d'autocorrélation de Moran

Contexte : L'autocorrélation spatiale désigne la ressemblance entre voisins. Deux objets ou pixels proches dans l'espace auront d'avantage de chances d'avoir des caractéristiques communes que s'ils sont éloignés. Toutefois cette propriété peut se révéler problématiques lorsque les entités en questions interviennent dans des analyses statistiques basées sur individus statistiques par définition non-dépendants les uns des autres. C'est pour cette raison que dans le cadre de notre travail nous devons garantir, autant que faire se peut, une autocorrélation nulle ou pour le moins minimale. Différents indice permettent d'évaluer cette autocorrélation entre des points, notamment les indices de Geary et de Moran qui donnent des informations relativement proches.

L'indice de diversité de Geary indique si la dissemblance entre points voisins est plus grande significativement qu'attendue d'un modèle aléatoire alors que **l'indice de Moran** indique si la ressemblance entre points voisins est plus grande significativement qu'attendu d'un modèle aléatoire. Il est généralement admis que l'indice de Moran est plutôt un indicateur global alors que l'indice de Geary est plus sensible aux petites différences de voisinage. Cela découle directement des formules de ces indices.

L'indice de Moran se calcule avec la formule suivante :

$$I = \frac{1}{N} \frac{\sum_i \sum_j W_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

Dans cette formule, N représente le nombre de cas, X_i la variable au pixel i, X_j la variable au pixel j, \bar{X} la moyenne de la variable X et W_{ij} une matrice de poids relatifs à la distance entre les pixels i et j (dite matrice de contigüité). La valeur de chaque point est comparé à la valeur moyenne de tous les points. L'indice de Moran se présente donc comme le rapport (pondéré par la distance entre points) de la covariance entre différents points sur la variance totale. L'indice de Moran varie entre -1 et 1 et sa valeur est proche de 0 lorsqu'il n'y a pas d'autocorrélation. Comme cela peut être fait pour un coefficient de corrélation, il est possible de tester la significativité de l'autocorrélation en comparant la distribution observée du I de Moran à une distribution théorique.

L'indice de Geary se calcule avec la formule suivante :

$$I_G = \frac{1}{2} \frac{\sum_i \sum_j W_{ij} (X_i - X_j)^2}{\sum_i (X_i - \bar{X})^2}$$

Dans ce cas, l'indice se présente comme le rapport (toujours pondéré par la distance entre points) de la variance entre points sur la variance totale. Dans cette formule, la valeur de chaque point est comparée individuellement à celle de tous les autres, ce qui explique sa plus grande pertinence au niveau local que l'indice de Moran. Il varie entre 0 et 2 et prend une

valeur de 1 en cas de non corrélation. Comme pour l'indice de Moran, sa significativité peut être testée par comparaison de la distribution de l'indice avec une distribution théorique.