

Modélisation économique des relations entre l'agriculture et l'environnement à l'échelle du bassin de la Seine

Février 2009

Pierre-Alain Jayet¹, Pierre Cantelaube¹, Paul Zakharov¹, Marie Génin¹,
Cyril Bourgeois¹, Audrey Polard¹, Céline Schott², Catherine Mignolet²
Florence Habets³, Pascal Viennot³, Elodie Philipe³, Otmane Souhar³

¹ INRA-Économie publique, 78850 Thiverval Grignon

² INRA ASTER- 88500 Mirecourt, France

³ Centre de Géosciences UMR Sisyphe 7619, 77305, Fontainebleau

1.	Partenaires	2
2.	Objectif.....	2
3.	Méthode générale	3
4.	Les fonctions dose-réponse	7
4.1.	Les rendements	7
4.2.	Les émissions de pollution.....	10
5.	La spatialisation.....	13
5.1.	Spatialisation des résultats du modèle AROPAj.....	13
5.1.1.	Localisation des cultures	14
5.1.2.	Localisation des Groupes Types.....	16
5.1.3.	Nouvelle unité spatiale	17
5.2.	Transfert des sorties AROPAj vers les USTI.....	18
5.3.	Utilisation des Petites Régions Agricoles	19
5.4.	Passage aux mailles MODCOU	21
6.	Résultats pour l'estimation des émissions de polluants azotés d'origine agricole.....	24
6.1.	Présentation des simulations test.....	24
6.2.	Exemple de spatialisation de résultats de simulation	25
6.3.	Interpolation sur les mailles MODCOU.....	25
6.4.	Effet « PRA » sur la spatialisation des résultats.....	26
6.5.	Premiers résultats sur les pertes en nitrate	27
7.	Eléments de discussion technique	29
7.1.	Avantages et limites de la méthode et de son application.....	29
7.1.1.	Informatique et calcul.....	29
7.1.2.	Résolution spatiale et frontières	30
7.1.3.	Validation des fonctions de réponse.....	31
7.2.	Questions ouvertes	31
7.3.	Poursuite du travail engagé.....	33
	Bibliographie.....	34
	Glossaire	34
	Indexation et notations.....	38
	Annexe : Groupes type V2-AROPAj des 8 régions concernant le Bassin de la Seine.....	39
	Région 121	39
	Région 131	40
	Région 132	43
	Région 133	44
	Région 134	46
	Région 135	48
	Région 136	50
	Région 151	53

1. Partenaires

Les partenaires impliqués dans les activités associés au thème « économie » et sur la durée du programme (2007-2011) sont les suivants :

1. *UMR-Economie Publique INRA-AgroParisTech - Gignon* :
Pierre-Alain Jayet, Pierre Cantelaube, Paul Zakharov,
Marie Génin (stage été 2008), Cyril Bourgeois (thèse 11/2008-11/2011),
Audrey Polard (thèse ULB, séjour de 3 mois)
2. *INRA - Mirecourt* :
Céline Schott, Catherine Mignolet
3. *Géosciences - ENSMP et Unité Sisyphe - Jussieu*:
Florence Habets, Pascal Viennot
Elodie Philipe (thèse 10/2007-09/2010), Otmane Souhar (postdoc)

2. Objectif

Globalement, il s'agit d'analyser, par la modélisation et sur le Bassin de la Seine, les impacts de différents scénarios économiques et/ou climatiques sur les activités agricoles et leurs effets sur le milieu. Les scénarios devraient intégrer l'évolution de l'environnement économique (Politique Agricole Commune, politique de promotion des cultures énergétiques, politiques environnementales) et de l'environnement physique (évolution du climat, avec l'adaptation des productions).

Les impacts sur les activités agricoles portent sur changement d'usage des sols agricoles aussi bien que sur les rendements. Les impacts sur l'environnement concernent dans un premier temps les pollutions azotées (N₂O, NO₃, NH₃). A terme, il s'agira d'intégrer dans l'analyse la régulation économique des pollutions d'origine agricole, en fonctions d'objectifs ciblés (réduction des émissions de gaz à effet de serre d'origine agricole sur la zone, maîtrise des concentrations en nitrate des eaux des aquifères du Bassin).

Trois modèles sont retenus pour l'analyse :

- le modèle économique d'offre agricole AROPAj ;
- le modèle agronomique STICS ;
- le modèle hydrogéologique MODCOU.

Ces modèles sont utilisés en forçage bilatéral.

L'objectif technique général est de coupler le modèle AROPAj et le modèle hydrologique de transfert MODCOU (Ecole des Mines – Géoscience et UPMC – SISYPHE), sous forme d'un couplage externe entre ces modèles. Le modèle STICS est utilisé pour la création de fonctions « dose-réponse » permettant de relier les activités agricoles et les émissions de polluants, fonctions qui sont nécessaires à l'association des deux autres modèles.

A court terme, pour l'année en cours, les objectifs sont :

1. estimer les relations entre les émissions de polluants et les activités agricoles, via des fonctions « dose-réponse » ;
2. estimer à différentes échelles spatiales, les vecteurs de pollution, en particulier les quantités de nitrate dont les transferts par lixiviation vont affecter les aquifères en relation avec le Bassin de la Seine ;
3. estimer les quantités de nitrate à la base racinaire selon le maillage du modèle MODCOU.

3. Méthode générale

On s'intéresse ici à la mise en œuvre de la modélisation autour d'un « scénario », scénario caractérisé par des éléments de contexte économique (PAC, régulation d'externalités, aides aux cultures énergétiques) et physique (climat) applicables aux systèmes d'exploitation agricole retenus dans le modèle économique, et pour une période de une ou plusieurs années sur un horizon prédéterminé. La méthode porte donc sur les calculs effectués autour du modèle AROPAj depuis l'identification et le traitement des paramètres affectés par le scénario (en entrée) jusqu'à la distribution spatiale des résultats de simulation à l'échelle d'une maille géographique compatible avec le modèle MODCOU (en sortie). Les éléments fournis par le modèle STICS sont supposés disponibles et fournis en amont (voir plus loin dans cette section ou dans les sections suivantes consacrées aux « fonctions de réponse »). Le bouclage intégrant une succession de simulations AROPAj et MODCOU, véritable finalité des travaux du PIREN-Seine sur le thème « économie », fera l'objet des travaux ultérieurs. La mesure des impacts d'un scénario affectant l'activité agricole jusqu'aux pollutions azotées suit le cheminement suivant :

- Identification des paramètres pertinents pour l'analyse (par exemple, révision du module « bilan azoté » du modèle AROPAj), et modification des paramètres du modèle AROPAj selon le scénario choisi, et, en cas d'un scénario dynamique, modification éventuelle des paramètres à l'itération $n+1$ en fonction des résultats de l'itération n ;
- Simulation réalisée (à chaque itération) à l'aide d'AROPAj, consistant à résoudre chacun des programmes d'optimisation associé à chacun des « groupes type d'exploitation agricole » de l'espace considéré : pour couvrir le Bassin de la Seine, on résout les programmes de l'ensemble des groupes type (GT) des 8 régions concernées ;
- Spatialisation des résultats pertinents pour l'appréciation du scénario (parmi les résultats fournis par AROPAj, qu'il s'agisse d'allocation des terres, de productions, de quantités de pollution agricole, ...) à l'échelle de la *maille élémentaire* ;
- Recombinaison des résultats à l'échelle de la maille d'intérêt pour le scénario étudié (par exemple la maille MODCOU).

Les émissions de polluants, à ce stade les polluants azotés (via la dénitrification et l'émission du GES : N_2O ; la lixiviation : NO_3 ; et la volatilisation : NH_3), sont déterminées en application de la méthode de couplage entre le modèle AROPAj et le modèle STICS. Il convient de distinguer deux niveaux dans l'élaboration des fonctions « dose-réponse ». Le 1^{er} niveau porte sur les fonctions de rendement, le 2nd sur les fonctions d'émission de pollution. Pour le 1^{er} niveau, la méthode a été initialement développée dans le cadre de la thèse de Caroline Godard (thèse soutenue le 11/12/2005), du programme FP6 GENEDEC, et de programmes GICC du Ministère de l'environnement (2002-2006). Elle a été étendue dans le cadre de la thèse d'Elodie Galko-Debove (extension à la France, avec l'automatisation des procédures de calcul). La section 4.1 en rappelle le principe. La méthode repose sur l'ajustement de « fonctions » obtenues par les résultats de simulations STICS appliquées à un ensemble de couples « précédent cultural – culture d'intérêt » et à des données météo compatibles avec l'observation économique.

Tout au long de ces travaux, ont été développées un ensemble de bases de données et d'applications automatisées utilisant le modèle STICS (ensemble dénommé ArTiX). Le 2nd niveau a été initié dans le cadre du programme européen GENEDEC et du programme ADD « DST » (Dégradation des sols par le tassement). Le principe en est rappelé en section 4.2.

Du point de vue de l'application économique (dans le modèle AROPAj), à l'échelle de la parcelle, la méthode de calcul des quantités de pollution est résumée par la Figure 1. Le raisonnement est le suivant : une fois choisie l'activité agricole (i.e. la culture), la

maximisation de la marge brute par unité de surface conduit à déterminer la quantité d'intrant et le rendement. Au niveau d'intrant ainsi déterminé, lui est associé un niveau de pollution émise. La modification des marges brutes unitaires peut se traduire par une modification de l'allocation des terres (voir la Figure 2). Dans le programme « AROPAj » de maximisation sous contraintes représenté Figure 2 en dimension 2, le changement des rendements et des charges en engrais par hectare se traduit par un changement des marges brutes par unité de surface de chacune des cultures, et donc par un changement du rapport des marges brutes unitaires entre elles. Un changement significatif de ce rapport peut impliquer un changement dans l'allocation des terres de l'exploitation agricole modélisée.

L'allocation optimale de la SAU entre les activités proposées relève du choix de l'agent économique représentatif (i.e. le groupe type AROPAj), et c'est la fonction du modèle AROPAj que de calculer la solution optimale. Aussi bien en termes de production que de pollution, il y a combinaison de deux effets pris en compte par le modèle économique : l'effet au niveau de la parcelle, et le changement d'usage des terres, le premier contribuant au second. La solution optimale dépend évidemment de l'ensemble des paramètres intégrés dans le modèle, et en particulier des paramètres d'environnement physique (i.e. le climat, tel que cela a été exploré par le travail de Galko-Debove, 2007) et économique (environnement caractérisé par la PAC, les prix, les instruments de régulation environnementale tels que les taxes et les quotas).

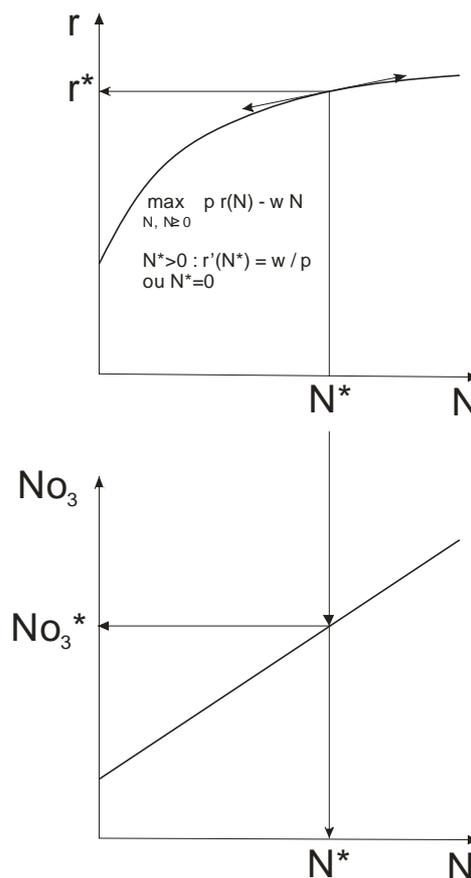


Figure 1. Consommation d'intrant, rendement, pollution NO₃, en fonction du rapport de prix w/p.

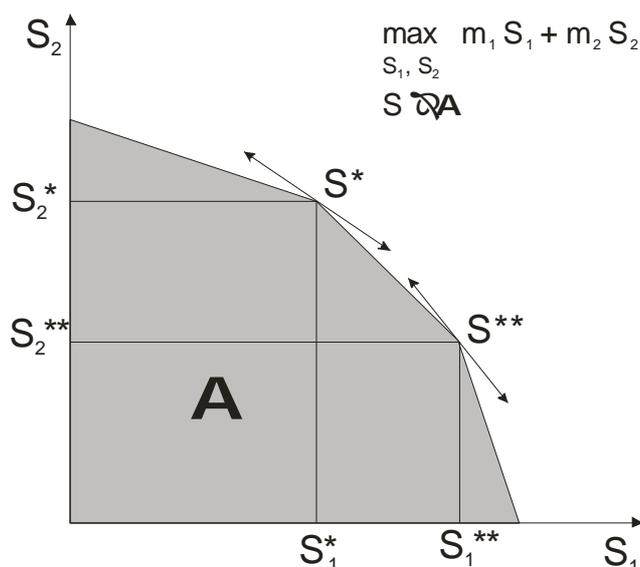


Figure 2. Exemple avec deux cultures (surfaces S_1 et S_2) du partage des terres d'une exploitation dépendant du rapport des marges brutes unitaires (m_1/m_2)

Associer l'activité agricole simulée par le modèle AROPAj et l'espace géographique constitue une étape importante de l'étude. La méthode de changement d'échelle des résultats fournis par AROPAj est fondée sur l'ensemble des informations géographiques susceptibles d'être disponibles pour l'Union Européenne, de façon *pérenne, homogène, et représentative*. Ces trois caractéristiques ont été mises en évidence lors du programme GENEDEC. Il convient de rappeler que l'unité géographique qui sous-tend la mise à disposition de données micro-économiques agricoles est la région « RICA » équivalente en France à la Région administrative. Une méthode de classification automatique a été élaborée pour définir les groupes type AROPAj, qui deviennent alors les *agents économiques représentatifs* de la production agricole. La classification repose sur :

- i. la typologie des exploitations agricoles selon la nomenclature du RICA (fondée principalement sur les parts de marge brute associées aux grandes catégories de productions de l'agriculture et de l'élevage) ; cette variable qualitative a été transformée en hiérarchisant les catégories RICA selon un axe [grandes cultures, hors sol], puis en leur conférant une valeur cardinale ;
- ii. l'altitude, répartie en 3 classes dans le RICA :]<300m], [300m,600m[,]>600[;
- iii. la « taille économique » répartie en 9 catégories dans le RICA, donnée en valeur normalisée.

Chaque groupe type AROPAj doit être associé à 15 individus au moins de l'échantillon du RICA (règle édictée par le service du RICA de la DG AGRI de la Commission Européenne pour le droit d'utiliser ces données individuelles).

La version de travail du modèle AROPAj dans cette phase du PIREN-Seine est la version V2 associée au RICA 2002 (UE-15), développée dans le cadre du programme GENEDEC. Le Bassin de la Seine interfère avec **54 groupes type** de la version V2-AROPAj répartis sur les 8 régions concernées (voir Figure 3).

Une nouvelle version du modèle AROPAj, V3 (RICA 2004, IE-25), a été « calibrée » en 2008, mais la base ArTiX n'a pas encore été alimentée et actualisée en conséquence.

La spatialisation proposée est fondée sur la méthode statistique développée dans le cadre du programme GENEDEC par Raja Chakir, économètre à l'UMR-EP, et Pierre Cantelaube en relation avec le JRC, Ispra. Elle mobilise les données physiques disponibles à l'échelle européenne (LUCAS, CLC, MARS, BSE, MNT). Une nouvelle application est ici proposée

en utilisant les données élaborées par l'équipe INRA de Mirecourt, à l'échelle des **petites régions agricoles (PRA)**.

L'organisation et la représentativité du RICA confèrent à la Région (en France) un rôle primordial, mais discutable quant à la capacité de bien représenter la diversité agricole. Les PRA ont été conçues sur la base de zones agricoles homogènes quant aux conditions pédo-climatiques et aux activités agricoles. L'adéquation entre activité agricole et espace géographique est importante pour l'évaluation et la régulation des externalités environnementales liées à cette activité et pour lesquelles la localisation est déterminante (i.e. les pertes d'azote en nitrate). Mais les données d'enquête recueillies sur les PRA sont difficiles à croiser avec les données du RICA, alors que le type de données micro-économiques présentes dans le RICA fait partie du socle fondateur du modèle AROPAj. Deux options émergent :

- (1) créer ou modifier la typologie AROPAj à partir des PRA ;
- (2) utiliser les PRA pour affiner la relation entre les groupes type AROPAj existant et l'espace géographique.

Compte tenu des données disponibles, de la difficulté de les reproduire dans la perspective d'actualiser les modèles, et des ressources humaines et financières disponibles, seule la seconde option paraît réaliste.

RICA - France

CODE REGION

121	Île de France
131	Champagne-Ardenne
132	Picardie
133	Haute-Normandie
134	Centre
135	Basse-Normandie
136	Bourgogne
141	Nord-Pas-de-Calais
151	Lorraine
152	Alsace
153	Franche-Comté
162	Pays de la Loire
163	Bretagne
164	Poitou-Charentes
182	Aquitaine
183	Midi-Pyrénées
184	Limousin
192	Rhône-Alpes
193	Auvergne
201	Languedoc-Roussillon
203	Provence-Alpes-Côte
204	Corse



Source: EC ESTAT-GISCO & AGRI-A.3
 Cartographie: EC GIS-AGRI
 Copyright: MEGRIN pour les bornes administratives

Figure 3. Régions RICA en France (en grisé, les Régions concernés par le Bassin de la Seine).

La mise en œuvre de la méthode d'ensemble suppose donc que soient réalisés les deux objectifs suivants :

- Etablir les fonctions « dose – réponse », pour un maximum de couples {*groupe type, culture*} préexistant dans le modèle AROPAj ;
- Localiser les groupes types AROPAj de façon probabiliste à l'échelle de la *maille élémentaire*.

Ces deux objectifs renvoient aux deux phases techniques et aux résultats exposés dans les deux sections suivantes. La dernière section développe des points critiques, des questions en suspens, et des suggestions en vue du forçage des résultats AROPAj-STICS vers le modèle MODCOU.

4. Les fonctions dose-réponse

Les 54 groupes type AROPAj contribuant à la couverture agricole du Bassin de la Seine et les 8 cultures présentes chez ces groupes types et pour lesquelles la méthode de couplage STICS-AROPAj a été validée conduisent à élaborer 329 séries de fonctions « dose-réponse » (rendement de production et émissions de pollution). Les cultures concernées sont :

- le blé dur ;
- le blé tendre ;
- l'orge (en réalité subdivisé en deux groupes pour STICS – orge d'hiver et orge de printemps – mais regroupées dans AROPAj du fait de l'agrégation RICA) ;
- le maïs grain
- la betterave à sucre ;
- la pomme de terre ;
- le colza ;
- le tournesol

Ces fonctions dose-réponse sont obtenus selon une procédure de traitement automatique mobilisant les applications ArTiX et le modèle STICS compilé en environnement UNIX sur le serveur de calcul de Grignon.

4.1. Les rendements

L'élaboration de la fonction de rendement, selon l'approche initiale de Godard, 2005 et Godard et al, 2008, est résumée par le processus suivant :

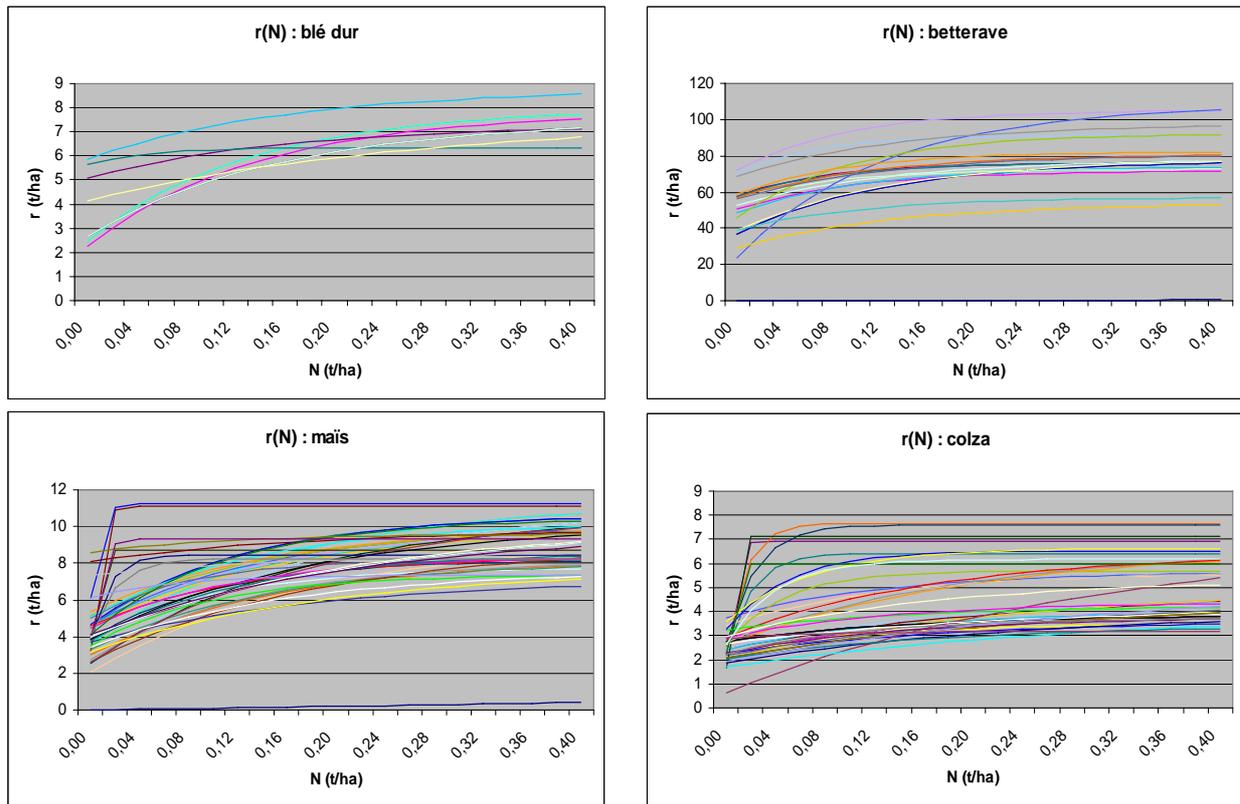
- 1) Pour la version du modèle AROPAj (i.e. V2-2002) : choix des « années climatiques » couvrant la succession { $n, n+1$ }, n étant l'année RICA (en réalité on utilise 3 années climatiques, compte tenu des périodes de recouvrement « année comptable » / « année climatique ») ; éventuel calcul de l'altitude moyenne des groupes type AROPAj (combinaison MNT et indicateur d'altitude RICA) ;
- 2) Pour la **Région** : étape préliminaire de mise à disposition des 5 sols « STICS » les plus représentatifs (à partir d'un travail réalisé par l'Unité INFOSOL de l'Inra-Orléans, dans le cadre de la thèse de Caroline Godard et du programme GENEDEC), identification des variétés pour les cultures préexistantes dans AROPAj ; recherche d'information sur les dates de semis ; recherche d'information sur les apports d'engrais azotés de synthèse (qualité, prix) ; hiérarchie des cultures irriguées et affectation des cultures effectivement irrigables compte tenu de l'estimation RICA-AROPAj de la surface irriguée ; estimation et répartition des effluents d'élevage sur les cultures, à partir des estimations d'effectifs animaux selon les données RICA-AROPAj ; ajustement éventuel
- 3) Pour un couple {*groupe type, culture*} préexistant dans le modèle AROPAj : constitution des fichiers STICS intégrant les 5 sols, 3 variétés ou dates de semis, 2 précédents possibles (soit 30 options) et, selon que la culture est « pivot » pour

- l'irrigation, passage de 30 à 60 options ; prise en compte de l'éventuel apport d'azote organique ;
- 4) Pour un jeu *{sol, variété / date de semis, précédent [, irrigation]}* : 31 simulations STICS pour un apport d'azote sous forme d'engrais de référence passant de 0 à 600 unités, par pas de 20 unités ;
 - 5) Ajustement de chacun des 30 (ou 60) séries de points calculés, selon la forme exponentielle $r(N) = B - (B - A) e^{(-\tau N)}$, qui est l'expression formelle retenue génériquement pour la fonction de rendement ;
 - 6) Présélection, parmi ces 30 fonctions ajustées, des fonctions telles que la valeur asymptotique B soit supérieure au rendement estimé R_0 selon les données RICA-AROPA_j ;
 - 7) Sélection de la fonction ajustée, parmi les fonctions présélectionnées, qui minimise la valeur absolue de l'écart entre la dérivée $r'(N_0)$ lorsque $r(N_0)=R_0$ et le rapport du prix de l'azote sur le prix du produit agricole ; on dispose alors, pour le couple *{groupe type, culture}*, des paramètres STICS associés *{sol, variété / date de semis, précédent [, irrigation]}* ;
 - 8) Mise à zéro des apports d'azote d'origine animale ; 31 simulations STICS pour un apport d'azote sous forme d'engrais de référence passant de 0 à 600 unités, par pas de 20 unités ; ré-ajustement de la courbe, qui deviendra **la fonction de rendement** du couple *{groupe type, culture}* ; mise au format « AROPA_j » des paramètres de la fonction de rendement.

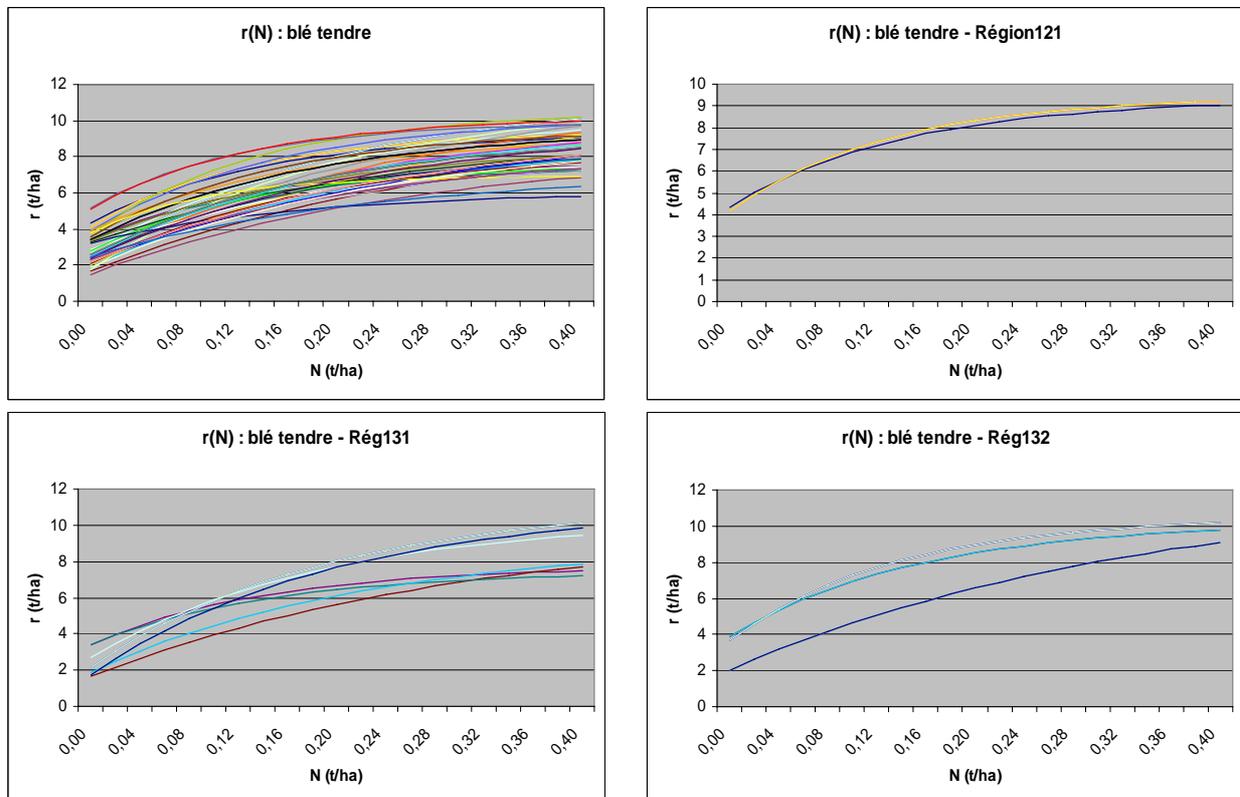
L'étape 8) ci-dessus n'est cependant pas réalisée lors de l'élaboration de grandes séries de fonctions de rendement, d'une part pour des raisons informatiques, d'autre part parce que cette étape ne s'est pas révélée déterminante dans les fonctions effectivement obtenues. La réalisation du processus qui conduit à inférer un jeu de paramètres STICS et à l'ajustement de la fonction de rendement associée à un couple *{groupe type, culture}* repose sur plus de 930 simulations STICS à pas de temps journalier sur 700 jours environ. Les données « sol » proviennent de la base européenne des sols, données « transformées » par la mise en œuvre de fonctions de pédo-transfert développées par l'Unité Infosol de l'Inra-Orléans. Les données météo proviennent du programme européen MARS.

Il est à noter que l'élaboration de fonctions complémentaires de rendement, indexées sur le type d'effluent d'élevage, a été amorcé lors du travail de thèse de Caroline Godard (2005), et qu'un module « épandage » a été incorporé au modèle AROPA_j. Cette option n'est pas encore validée. Par ailleurs, durant le travail de thèse d'Elodie Galko-Debove (2007), les fonctions de rendement ont été ajustées selon différents scénarios climatiques (provenant du modèle ARPEGE-Climat), avec ou sans forçage CO₂. Les paramètres physiques autres que le climat sont pour partie inchangés (i.e. le sol), tandis que d'autres pourraient l'être (date de semis, variété, irrigation) selon le type de scénario retenu.

La variabilité obtenue (voir les Figures 4 et Figures 5) reflète bien évidemment la diversité des conditions pédo-climatiques (plus précisément celle des sols) et, en association avec celles-ci, la diversité des itinéraires techniques (variétés, dates de semis, précédents culturels). Il faut également compter avec les aléas des pratiques et des comportements pour tenter d'expliquer la diversité au sein même des Régions (que l'on peut observer par exemple au niveau du rendement de référence (R_0) utilisé pour la sélection et l'ajustement des fonctions), même quand elles présentent une relative homogénéité du point de vue des conditions de production et des activités agricoles choisies (voir les Figures 5).



Figures 4. Fonctions de rendement pour différents couples {groupe type, culture} : blé dur, betterave, maïs, colza pour l'ensemble des GTs du Bassin de la Seine.



Figures 5. Fonctions de rendement du blé tendre, pour les groupes types des Régions couvrant le Bassin de la Seine, et pour les Régions 121 (Ile-de-France), 131 (Champagne-Ardennes) et 132 (Picardie).

Il conviendra cependant de revenir sur l'amplitude du spectre de valeurs associé à certains paramètres, par exemple la valeur à l'origine (le paramètre A) de la fonction de rendement du blé dur, ou encore la valeur asymptotique (i.e. le potentiel B) du colza. La méthode d'ajustement des fonctions de rendement à destination du modèle AROPAj est certes fondée sur le rendement et le système de prix observés (estimations fondées sur le RICA à l'échelle des groupes type AROPAj). Le développement de nouvelles versions du modèle STICS est susceptible de fournir des résultats « recalibrés » en regard de certaines productions végétales. Mais, dans le « scénario référence », les résultats « AROPAj » en matière de rendement, de charge, et d'émission de polluants devrait en être peu affecté, du fait de la procédure d'ajustement des fonctions sur la base des observations précitées. Par contre, l'effet d'un changement de fonctions de réponse, combiné au changement de scénario pourrait être significatif.

4.2. Les émissions de pollution

L'élaboration d'une fonction de rendement pour un couple $\{groupe\ type, culture\}$ repose donc sur la sélection d'un jeu de paramètres physiques en entrée du modèle STICS permettant de restituer un rendement et une « tangente » proches de l'observation. Une fois ces paramètres sélectionnés, on dispose avec le modèle STICS d'un ensemble de « sorties » utiles pour notre analyse. On ne discute pas ici de la calibration et de la validation des modules qui rendent ces sorties disponibles.

Les variables d'intérêt en sortie du modèle STICS sont :

- i. Le protoxyde d'azote - N_2O - (dénitrification) ;
- ii. Le nitrate - NO_3 - (lixiviation) ;
- iii. L'ammoniac - NH_3 - (volatilisation).

Aux 31 points permettant d'ajuster définitivement la courbe de rendement (une fois les paramètres STICS déterminés), on peut associer 3 séries de 31 points représentant l'évolution de chacune des émissions (N_2O , NO_3 , NH_3) lorsque varie l'apport d'azote. Le choix d'une forme fonctionnelle a été guidé par la simplicité, même si elle a été discutée (pour N_2O , voir Durandau et, 2008).

Comme la Figure 1 le suggère, nous avons retenu la fonction affine $e(N) = A N + B$. Il s'agit alors d'estimer 2 paramètres pour chaque triplet $\{groupe\ type, culture, pollution\ azotée\}$. Dans un premier temps, les estimations sont effectuées séparément (31 points), mais la question de l'estimation jointe des paramètres et de l'étude de la significativité des contrastes (entre cultures, entre régions, ...) est en suspens.

L'interprétation que l'on peut donner aux paramètres de type A (le facteur d'émission) et de type B (la « constante ») renvoie à des questions ouvertes soulevées dans la section 6.5. On relèvera simplement à ce stade que l'indexation de la « constante » par la culture et/ou le groupe type pose la question des émissions « naturelles » (en l'absence d'apport d'azote par l'agriculteur). Pour le calcul des émissions de N_2O d'origine anthropique, rappelons que l'IPCC ne retient que le facteur d'émission A , qui plus est sans l'indexer. Néanmoins, sans même tenir compte de l'apport d'azote, le choix de la culture est d'origine anthropique, et ignorer la constante B dans le calcul des émissions est alors contestable. Enfin, pour un type donné de pollution, quand bien même la constante B serait « universelle » et sans rapport avec les choix humains, on ne peut l'ignorer lorsqu'il s'agit d'évaluer le niveau total de pollution. Ainsi en est-il de la concentration en nitrate d'une masse d'eau. Le point important est alors le niveau de pollution de référence (parfois confondu avec le niveau de pollution souhaitable ou

« optimal », et qui de toutes façons n'a pas de raison d'être égal à zéro), lorsque seul compte l'effet additionnel d'origine anthropique comme c'est le cas avec les émissions de gaz à effet de serre.

Les résultats obtenus pour chacun des 3 types de pollution et parmi les 301 premiers couples {*groupe type, culture*} étudiés (hors betterave) sont de nature assez différente d'un type de pollution à l'autre, mais également d'une culture à l'autre. Tout d'abord, le modèle STICS peut conduire à l'absence de perte d'azote sous l'une ou l'autre forme (N₂O, NO₃, NH₃), voire pour les 3 formes. C'est le cas chez plusieurs groupes types pour le colza, pour le tournesol et le maïs (principalement en matière de perte NO₃, également NH₃), pour la betterave (NO₃ et NH₃), et pour la pomme de terre (NH₃). Les « constantes » d'émission N₂O pour les blés et les orges sont en général assez proches, contrastant fortement avec les valeurs obtenues pour les autres cultures. La plupart des facteurs d'émission pour N₂O sont estimés à des niveaux compris entre 0.015 et 0.025 (donc sensiblement plus forte que la valeur retenue par l'IPCC qui est de 0.010).

Les résultats les moins attendus concernent les pertes NO₃, souvent estimées à la valeur dans 146 cas pour le facteur d'émission *A* et dans 106 cas pour la constante *B*, avec une concentration relative sur certaines Régions (Ile-de-France par exemple). Le Tableau 1 indique la fréquence d'estimation de valeur nulle pour les paramètres d'émission pour les 301 couples {*groupe type, culture*} possibles (hors betterave) sur le Bassin de la Seine.

En marge de l'établissement de ces fonctions, il conviendra d'intégrer au modèle AROPAj un facteur d'émission à valeur par défaut, pour les cultures pour lesquelles il n'y a pas de fonction dose-réponse associée (certaines cultures de vente, les fourrages et prairies). Cela joue essentiellement pour les pertes d'azote sous forme NH₃ et NO₃ (pour N₂O, on dispose du facteur d'émission proposé par l'IPCC).

	N ₂ O	NH ₃	NO ₃	Tous
<i>A</i>	5	49	146	5
<i>B</i>	34	62	106	11
<i>A et B</i>	5	49	101	5

Tableau 1. Nombre de cas où les paramètres d'émission de pollution *A* et *B* de la relation $e(N) = A N + B$ sont estimés à 0 parmi 301 couples AROPAj {*groupe type, culture*} affectant le Bassin de la Seine.

N ₂ O	bd	bl	oh	ma	bt	pt	cz	tr
121								
131							4	
132								
133								
134								
135								
136								
151							1	

Tableau 2. Nombres de cas où le facteur d'émission « A-N₂O » est estimé à 0.

NH3	bd	bl	oh	ma	bt	pt	cz	tr
121			1		1			
131		4	1	8	5		4	4
132								
133								
134	1	5	1	3				
135			4	3			2	
136								
151							1	1

Tableau 3. Nombres de cas où le facteur d'émission « A-NH3 » est estimé à 0.

NO3	bd	bl	oh	ma	bt	pt	cz	tr
121	3		3	3	3	1		2
131		7	1	8	5		4	4
132		6	5	8	8			1
133		4			1			
134	2	7	3	5	3	1		2
135		1	2				1	
136		8	5	1	3			9
151		3	1	4	1		2	2

Tableau 4. Nombres de cas où le facteur d'émission « A-NO3 » est estimé à 0.

La Figure 6 donne un exemple de « sortie » STICS utilisée pour élaborer les fonctions de réponse en production et pollution.

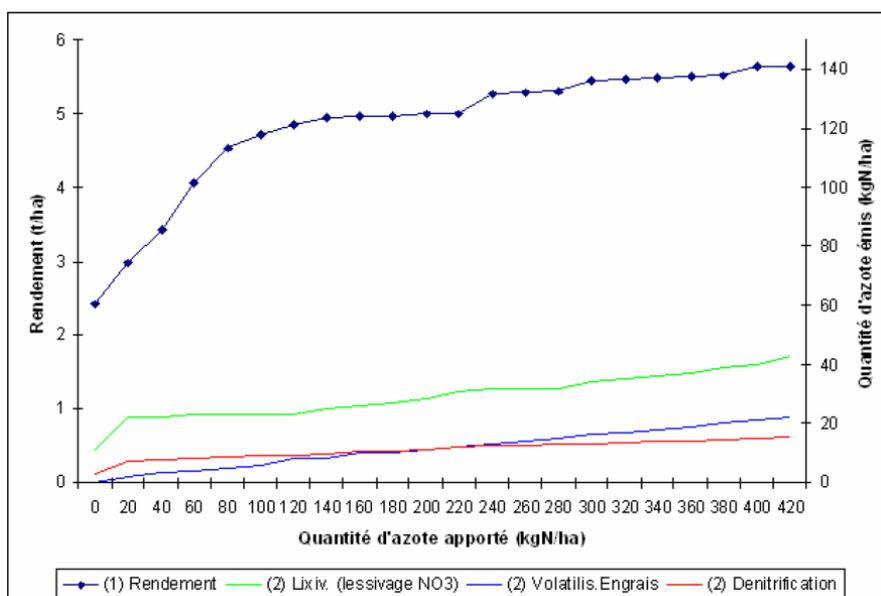


Figure 6. Réponse en production et pollution à l'apport d'azote : exemple de l'orge chez le groupe type 7 (région 131).

5. La spatialisation

Cette section repose sur différents modes d'utilisation, de transformation, et de production d'information spatialisée sous forme probabiliste, et particulièrement :

- a. Calcul de la probabilité a priori de la présence d'une activité agricole de type j sur une cellule l (maille CLC) selon une relation expliquant la présence observée (LUCAS) à partir de données physiques (sol, climat, altitude, ...)
- b. Calcul de la probabilité de présence d'une activité agricole de type j sur une cellule l qui permette de rapprocher au mieux les surfaces moyennes ainsi obtenues et les surfaces estimées par le RICA (à l'échelle régionale)
- c. Calcul de la probabilité de présence d'un groupe type AROPAj k sur une cellule l (la mise à disposition du RICA interdisant toute utilisation directe, et a fortiori toute localisation d'information à caractère individuel, les groupes type AROPAj sont des entités virtuelles représentatives que nous cherchons à distribuer dans l'espace à partir des seules informations disponibles autorisées pour l'utilisation statistique)
- d. Distribution de l'allocation des terres sur l'ensemble des mailles élémentaires (l) à l'échelle pertinente pour le RICA (la Région), en utilisant la répartition de la SAU selon les groupes type AROPAj (k) et l'estimation initiale des surfaces consacrées aux différentes activités agricoles (j) ; cette distribution est dégradée dès lors que l'on utilise la solution calibrée AROPAj à la place de l'allocation des terres initialement et directement estimée à partir du RICA ; plus généralement, tout résultat AROPAj (surfaces, productions en volume ou valeur, pollutions ...) peut être distribué sur l'ensemble des mailles élémentaires ; avec l'hypothèse de groupes type AROPAj structurellement inchangés (effectif représenté, SAU, quotas, capital animal), la méthode s'applique directement pour n'importe quel scénario proposé ;
- e. Amélioration de la méthode (dès l'étape b ci-dessus) en cas de mise à disposition d'information complémentaire sur l'utilisation des terres ; ce sera le cas avec des données sur l'allocation des terres à l'échelle des Petites Régions Agricoles ;
- f. Regroupement des mailles élémentaires identiques quant à leur « contenu » par le SIG utilisé ;
- g. Interpolation à partir du contenu des polygones précédents, au prorata des intersections de surface, lorsqu'un nouveau maillage est introduit pour l'analyse (ce qui sera le cas avec le couplage AROPAj-MODCOU).

5.1. Spatialisation des résultats du modèle AROPAj

Les méthodes et objectifs présentés ci-dessous reprennent la procédure établie lors du projet européen GENEDEC (WP3.4 *Distribution Spatiale des Groupes-types AROPAj*), par l'INRA (UMR Economie Publique, Grignon), la société EnviCare et le JRC (EC-JRC, unité Agri-Env de l'Institut de l'Environnement Durable).

Connaissant les principales orientations technico-économiques des exploitations agricoles présentes au sein des régions RICA, et leurs regroupements en groupe-types agro-économiques, l'objectif est d'élaborer un formalisme de localisation géographique des groupes types, sur la base de données biophysiques et/ou statistiques disponibles en Europe.

Le résultat est un modèle spatial implémenté dans un Système d'Information Géographique, qui permet la réalisation d'une cartographie de la localisation des GT au sein des régions RICA, à une résolution cohérente avec les données d'entrée. Ainsi la simulation de l'impact

des politiques économiques (par exemple différentes options de réforme de la PAC) sur les groupes type peut être traduite dans l'espace géographique à différents niveaux de résolution. On peut donc « produire » une représentation spatiale des impacts, de la même façon que l'on peut produire les impacts sur les agents économiques utilisateurs de cet espace. Les impacts spatialisés peuvent être produits comme le sont les impacts associés à toutes les variables du modèle AROPAj (marge brute moyenne à l'hectare, valeur marginale du facteur « terre agricole », valeurs marginales associées aux autres facteurs quasi-fixes, allocation des terres entre les différentes activités végétales, production et pollutions).

La spatialisation proposée permet de s'affranchir de l'unité régionale sur la base de laquelle sont définis les groupes type statistiquement représentatifs de l'activité agricole régionale. L'intérêt va au-delà de la capacité d'illustration géographique de résultats de modèle (utile au demeurant pour l'appréciation des politiques publiques), en particulier lorsque l'analyse d'un problème nécessite le couplage de modèles à un fort niveau de résolution spatial (i.e. dans le cas de l'analyse physique et économique de la pollution des nappes d'eau en nitrate d'origine agricole).

La méthode de spatialisation a été développée pour être appliquée à l'ensemble du territoire de l'Union Européenne, les principales contraintes étant d'établir un formalisme homogène à ce niveau, et de découpler, lors de la modélisation géo-statistique, les variables biophysiques « stables » (caractéristiques des sols), les variables agro-économiques à disponibilité intermittente (utilisation et occupation des sols), et les variables économiques agrégées (statistiques agricoles de l'Union Européenne). Les limitations techniques des données à utiliser pour cette étude sont l'homogénéité et l'exhaustivité sur le territoire européen.

Il s'agit de discriminer spatialement à un niveau plus fin les groupes types AROPAj définis au départ sur une base régionale (les régions RICA) et selon les classes altitudinales (les GT regroupant des exploitations selon 1, 2 ou 3 parmi les 3 classes d'altitudes distinctes reconnues par le RICA). Cette discrimination est réalisée sur la base des données d'occupation du sol (CORINE Land Cover, LUCAS) et des variables dérivées des bases de données « sol » et « climat » représentant les aptitudes du milieu pour les productions associées aux différents groupes types.

Pour cela le processus de spatialisation suit 2 étapes :

- Spatialisation de l'utilisation du sol à une échelle fine (« maille » ou « cellule » élémentaire) par le calcul de la probabilité de rencontrer, sur chaque maille, tel type d'activité agricole ;
- Spatialisation des groupe-types AROPAj par le calcul des contributions de chacun des GTS à l'activité agricole sur chacune des mailles élémentaires dans une Région donnée.

5.1.1. Localisation des cultures

La « carte » produite à la fin de cette étape a pour support la base de données géo-référencées CORINE Land Cover (CLC) (CEC, 1993 et EEA, 2005), carte de l'occupation des sols qui recouvre l'ensemble du territoire de l'Union Européenne (EU25) par cellule de 100mx100m. CLC qui est une carte thématique affectant à chaque cellule une classe d'occupation du sol, se révèle donc être un outil très précieux. Mais la nomenclature qui lui est attachée ne permet pas de distinguer correctement les diverses activités agricoles. Il convient d'utiliser d'autres données pour affiner l'information fournie par CLC et de rechercher une méthode appropriée de calcul de la probabilité d'occupation des sols pour chaque cellule au sein de chaque région qui valorise ces données.

La méthode est la suivante :

- On dispose de l'enquête statistique LUCAS comportant une nomenclature riche de nombreuses classes relatives aux pratiques agricoles (Gallego, 2002). On dispose ainsi d'une spécification détaillée du type de culture qui peut être rencontré localement.

LUCAS fournit un échantillonnage du territoire grâce à un ensemble de points d'observations géo-référencés, sans pour autant constituer une carte ;

- Un modèle est construit pour établir des relations entre les classes LUCAS et des variables explicatives comportant :
 - o les classes CLC,
 - o les propriétés du sol (carte des sols INRA/EUROSOIL),
 - o l'altitude et la pente (modèle numérique de terrain),
 - o les précipitations et les températures (base de données climatiques MARS)¹ ;
- Ce modèle est calibré sur les points LUCAS d'une région RICA puis appliqué sur l'ensemble des cellules 100mx100m de cette région. Il en résulte une carte des probabilités des cultures (usage du sol) pour les différentes régions RICA, ceci à l'échelle du pixel (100x100m, soit 1ha). **Chacune des mailles élémentaires est affectée à une et une seule Région.**

Cette méthode qui repose sur la superposition cartographique de différentes sources de données et par la mise en œuvre d'un modèle économétrique d'ajustement est illustrée par la Figure 7.

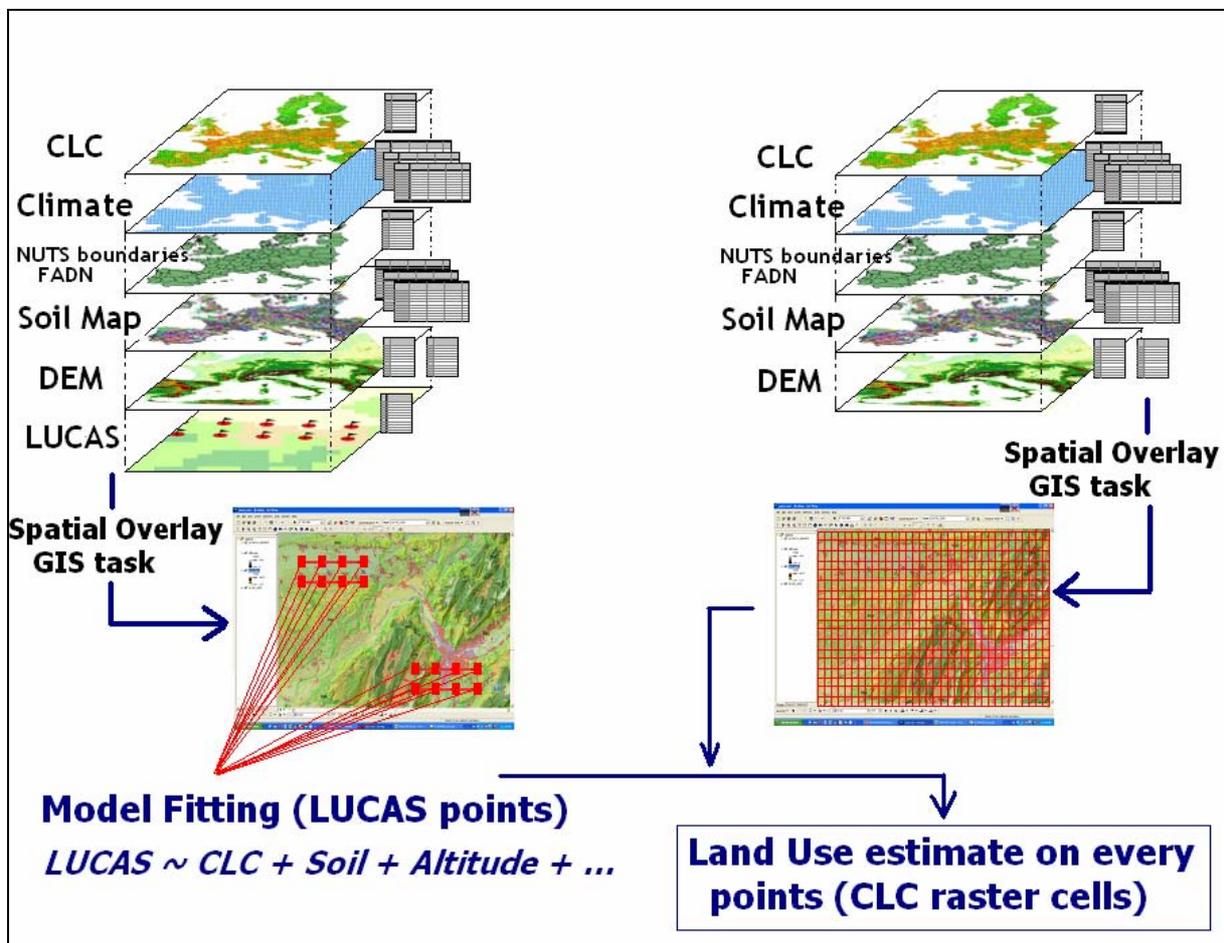


Figure 7 – méthode d'estimation de la carte d'occupation des sols

¹ A noter que les données d'entrée de cette méthode nécessitent une combinaison spatiale de plusieurs « couches » d'information (intégrant CLC, support de la carte). Ce croisement (« overlay ») de données spatiales est effectué grâce à des Systèmes d'Information Géographique (SIG).

Le modèle retenu suit ce que propose Chakir (2008), suivant une stratégie à 2 étapes :

- d'abord une modélisation « *Logit Multinomial* » pour l'estimation des paramètres d'une relation formelle caractérisant la probabilité de « trouver » une activité « LUCAS » conditionnellement aux valeurs des variables explicatives,
- puis une méthode d'optimisation de type « *Generalized Cross Entropy* » qui recalculent les probabilités de sorte que les surfaces régionales occupées par les différents types de culture soient égales aux surfaces estimées à l'aide des données du RICA.
- En 1^{ère} étape, le modèle Multinomial (MNL) permet d'attribuer à chaque unité spatiale i la probabilité d'accueillir une catégorie « LUCAS » j de couverture du sol. La distribution des « points LUCAS » sur les cellules CLC conduit à définir l'observation y_{ij} indiquant la présence de l'activité j sur la cellule i (y_{ij} à valeur 0 ou 1). L'observation dichotomique conduit à proposer un modèle probabiliste qui « explique » l'apparition d'un événement j conditionnellement aux caractères de la cellule i : $\tilde{p}_{ij} = F(x_i; \beta_j)$. On y exprime la présence de l'activité j (sur l'intervalle $[0,1]$) selon les caractères physiques x_i (CLC, sol, climat, altitude, ...) par une fonction de probabilité de la forme choisie :

$$\tilde{p}_{ij} = \frac{e^{x_i \beta_j}}{\sum_h e^{x_i \beta_h}} \text{ avec } \forall j : \sum_i \tilde{p}_{ij} = 1 \text{ et } \forall i, j : \tilde{p}_{ij} \geq 0$$

Cette étape est conduite pour chacune des Régions, ce qui signifie que le vecteur des paramètres β dépend de la Région. L'estimation de β conduit à l'estimation des probabilités a priori \tilde{p}_{ij} .

- En 2^{ème} étape, le modèle « cross entropy » (GCE) vise à déterminer les niveaux de probabilité p_{ij} les plus proches des probabilités a priori compatibles avec les niveaux de surface que l'on observe et utilisées pour le modèle économique. Ces dernières sont estimées à partir du RICA (base pour AROPAj). Ces nouvelles probabilités sont calculées (région par région) en résolvant le problème d'optimisation non linéaire (« minimisation de l'entropie croisée ») :

$$\min_{p_{ij}} CE = \sum_i \sum_j p_{ij} \ln \frac{p_{ij}}{\tilde{p}_{ij}}$$

$$\forall i : \sum_j p_{ij} = 1; \forall j : \sum_i p_{ij} u_i = R_j$$

Dans ce programme, u_i désigne l'aire de la cellule i (a priori 1 ha, éventuellement diminuée de la part non agricole estimée par CLC) et R_j est la surface régionale occupée par la classe de couverture du sol j .

5.1.2. Localisation des Groupes Types

Disposant de la carte d'occupation des sols, il s'agit de localiser les groupes types du modèle AROPAj présents dans une région RICA donnée. Rappelons que la typologie AROPAj ne repose que sur des critères renseignés par le RICA (OTEX, classe d'altitude, dimension économique normalisée). On peut affiner la probabilité de trouver le groupe type k sur la cellule i en déterminant cette probabilité conditionnellement à la classe d'altitudes a du groupe type (celui-ci pouvant regrouper des individus d'altitudes différentes). Rappelons enfin que la localisation des activités dépend des classes d'activités effectivement reconnues par LUCAS. Ces classes d'activité diffèrent a priori des activités reconnues par le RICA (par exemple LUCAS faisant références aux « céréales » alors que le RICA différencie blé tendre, blé dur, orge, maïs, ...). Les classes d'altitude sont évidemment aisées à spatialiser, du fait de la disponibilité des modèles numériques de terrain (MNT).

A partir de l'affectation des groupes types par classe d'altitude pour la région considérée, on peut déterminer la surface destinée par le groupe type k à l'activité j pour la classe altitudinale

a , notée $S_{ik|a}$. Connaissant, à partir de la localisation des activités fournie en section 5.1.1, les probabilité p_{ij} de trouver les type d'activité LUCAS j sur la cellule i (dont l'appartenance à la classe d'altitude a est connue par le géo-référencement), on peut calculer :

$$i \in a : q_{ik} = \sum_j p_{ij} \frac{S_{ik|a}}{\sum_h S_{ih|a}} \quad (1)$$

On vérifie aisément, pour la région considérée, que :

$$\forall i : \sum_k q_{ik} = 1$$

Ainsi, q_{ik} s'interprète comme la probabilité de trouver le groupe type k sur la cellule i ou encore comme sa contribution à l'activité agricole de cette cellule.

On doit enfin tenir compte du fait qu'une maille CLC puisse ne pas être totalement « agricole ». CLC nous fournit la part « NAU » (*non agricultural use*). On peut alors recalculer la probabilité de localisation des groupes type rapportée à la surface totale. On note q_{i0} la part de surface NAU sur la cellule i . Pour ne pas multiplier les notations, on conserve q_{ik} pour désigner la probabilité de localisation du groupe k pondérée par la part de surface agricole. On substitue en quelque sorte $q_{ik}(1-q_{i0})$ à q_{ik} et en associant le « groupe NAU » aux K groupes types AROPAj de la Région on a :

$$\forall i : \sum_{k=0,K} q_{ik} = 1$$

En considérant que les cellules élémentaires sont de même dimension, normalisée à 1 (unité de surface), on retrouve la SAU du groupe k telle qu'elle est estimée sur toute la Région par le SIG :

$$\forall k : \sum_i q_{ik} = SAU_k$$

Il convient de noter que la surface SAU_k est a priori différente de la SAU du groupe k utilisée dans AROPAj. En effet, le modèle économique ne représente qu'une partie de l'agriculture régionale, du fait des biais du RICA ou du fait qu'AROPAJ n'intègre pas toutes les activités agricoles. La technique adoptée consiste en premier lieu à placer en NAU les surfaces identifiées par CLC qui concernent des activités explicitement non prises en compte par AROPAj (i.e. la vigne, l'arboriculture, le maraichage). En second lieu, le mode de calcul des probabilités de localisation des groupes type tient compte du biais RICA (voir la relation (1)). On peut alors redistribuer toute « variable AROPAj » dans l'espace géographique de l'étude, à la résolution de la cellule élémentaire, après avoir simplifié la géostatistique du problème en manipulant des agrégats homogènes de cellules élémentaires.

5.1.3. Nouvelle unité spatiale

En termes d'unité spatiale, le résultat de la superposition et de l'intersection des différentes couches d'information est une carte formée de polygones. Ces polygones regroupent les cellules CLC ayant les mêmes propriétés (même classe CLC, même type de sol, même données météorologiques...). La Figure 8 illustre ces nouvelles unités spatiales (l'exemple porte sur une partie de la région Haute-Normandie). **Par construction, ces polygones forment une partition de chacune des Régions**, et sont dénommés unités spatiales de transformation de l'information (USTI).

Les traitements géostatistiques sont alors effectués sur un nombre de polygones bien inférieur au nombre de cellules élémentaires. Par ailleurs, ces traitements ne sont pas appliqués à toutes les USTI. Ainsi le modèle MNL n'est pas appliqué sur les zones « clairement » non agricoles : zones urbaines ou cours d'eau (cette information est fournie par CLC). Sur la

Figure 8 on distingue par exemple les zones urbaines de Rouen (à l'est) et du Havre (à l'ouest) et la Seine, zones géographiques non traitées par la modélisation.

A ce stade, l'important est de considérer les USTI comme de nouvelles unités spatiales s'imposant pour les différentes étapes de l'économétrie spatiale (MNL et CE). Ce seront également les unités élémentaires de calcul pour la distribution spatiale des « sorties AROPAj », quelles que soient les caractéristiques des couches supplémentaires d'information éventuelle utilisées (voire les PRA en section 5.3).

La surface agricole d'une USTI est notée Z_s :

$$\forall s : \sum_{i \in S} q_{ik} = Z_s$$

Il convient enfin de noter que les USTI ne forment pas des zones contiguës.

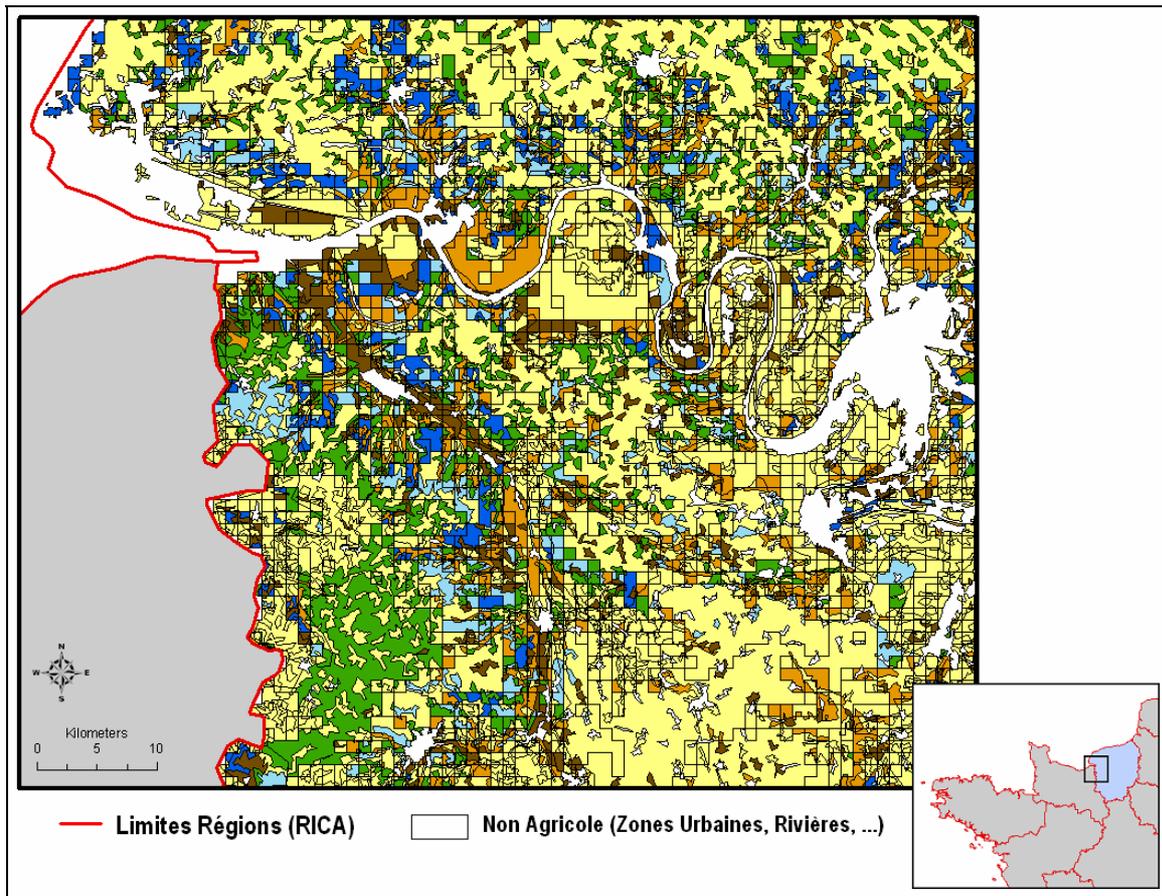


Figure 8 – Exemple de la nouvelle carte d'occupation du sol et des nouvelles unités spatiales ; les des polygones formées par les cellules 100mx100m CORINE LC (Couleurs et données fictives).

5.2. Transfert des sorties AROPAj vers les USTI

Les USTI sont les ensembles de cellules spatiales élémentaires « homogènes » définies pour permettre la distribution (i.e. le transfert) des « sorties » du modèle AROPAj sur un espace géographique de travail. Rappelons que les USTI forment une partition de chacune des régions à la base de la formation des groupes type AROPAj (elles peuvent être déterminées pour chacune des régions prises en compte par une version donnée du modèle AROPA). La fonction élémentaire de transfert est la fonction de probabilité de localisation des groupes type AROPAj. Par construction des USTI, la probabilité de localisation d'un groupe type AROPAj

sur une USTI est égale à la probabilité de localisation de ce groupe type sur toutes les mailles élémentaires qui la composent.

On conserve la notation q_{sk} pour désigner la probabilité de localisation du groupe type k sur l'USTI s . Rappelons que l'on a noté Z_s la SAU de l'USTI s .

Considérons une variable de « sortie » AROPAj, notée $Y_{.k}$ pour le groupe type k . On suppose que la sortie AROPAj est évaluée par unité de SAU. Pour toute USTI s « agricole » ($Z_s > 0$, i.e. $q_{s0} < 1$) et donc pour toute cellule élémentaire qui la compose, la variable Y est distribuée sur l'espace géographique par la relation :

$$\forall s : Y_{s.} = \sum_{k>0} \frac{q_{sk}}{1 - q_{s0}} Y_{.k}$$

L'information distribuée est évaluée par unité de SAU sur la zone (et non par unité de surface intégrant la surface NAU).

La localisation des groupes type est un moyen de transférer n'importe quelle « sortie » AROPAj sur un espace géographique donné. Les USTI sont définies pour faciliter ce transfert. Elles seront utilisées pour le transfert de l'information (i.e. les sorties AROPAj) depuis le maillage USTI vers n'importe quel autre maillage (i.e. le maillage MODCOU, voir la section 5.4).

5.3. Utilisation des Petites Régions Agricoles

Par construction, le modèle AROPAj induit de l'homogénéité à l'échelle régionale, puisque la représentativité des données micro-économiques servant à le calibrer est réputée vérifiée à cette échelle. La discrimination entre groupes types selon l'OTEX, l'altitude et la dimension économique peuvent mal émerger, en particulier lorsque l'échantillon enquêté est de taille réduite. La question s'est posée d'utiliser de l'information complémentaire (dérogant aux principes d'existence à l'échelle européenne et de reproductibilité dans le temps énoncés pour le développement d'AROPAJ).

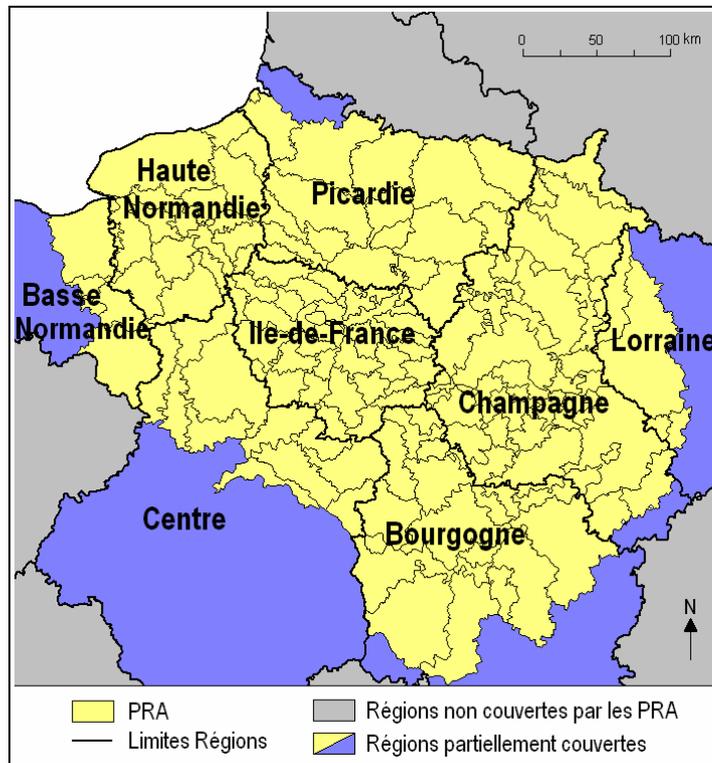
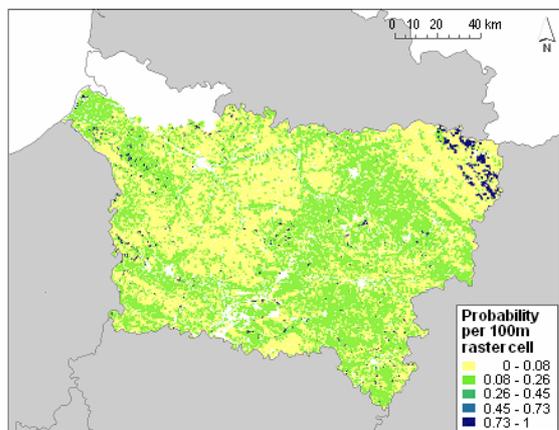


Figure 9 – Situation géographiques des PRA et régions administratives

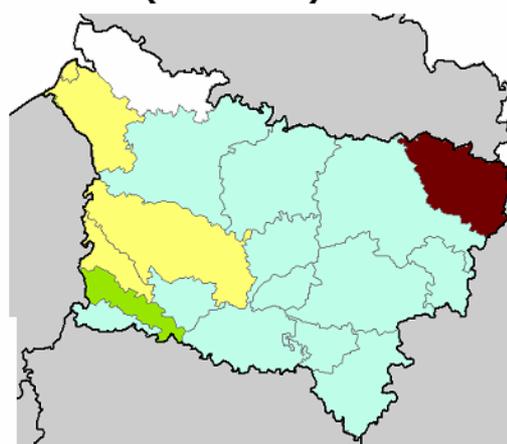
Exemple de re-calibration de la Carte de Couverture des Sols sur les PRA

Région Picardie (RICA 132)

1 Estimation de la classe de couverture du sol "Prairie Permanente", calibrée sur la région



2 Evaluation de cette carte vs observations (surfaces) sur les PRA



3 Carte des "Prairies Permanentes" après la calibration sur les PRA

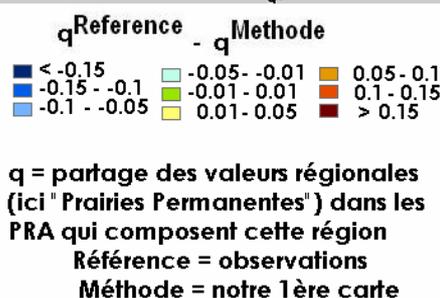
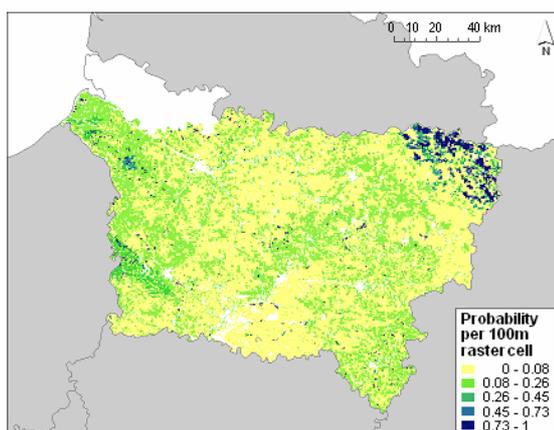


Figure 10 – Exemple illustrant la re-calibration de la carte de couverture du sol sur la région Picardie : probabilité de localisation des prairies permanentes respectivement fondée sur l'allocation des terres connues à l'échelle de la Région (1) et des PRA (3), et probabilité moyenne par PRA (3).

Les PRA sont des zones contiguës, à la différence des USTI. L'utilisation de l'information PRA vient affiner le calcul des probabilités de localisation des activités agricoles, et par conséquent les probabilités de localisation des groupes types. Cette information intervient au stade de la modélisation économétrique *CE*, sans modifier le stade *MNL*.

Le partenaire PIREN *INRA Mirecourt* a fourni les données des surfaces cultivées (en 2000) sur les Petites Régions Agricoles² (PRA) composant le bassin de la Seine. Elles ont l'avantage de s'inscrire dans les Unités administratives existantes (formant en France une

² Les régions agricoles et petites régions agricoles ont été définies (en 1946 et 1949, par l'INSEE) pour mettre en évidence des zones agricoles homogènes. La Région Agricole (RA) couvre un nombre entier de communes formant une zone d'agriculture homogène. La Petite Région Agricole (PRA) est constituée par le croisement du département et de la RA. Par définition, dans une Petite Région Agricole on retrouve une même vocation agricole dominante.

partition des Départements, et donc des régions). La Figure 9 illustre la superposition du Bassin de la Seine, des PRA, des 8 régions concernées. Les PRA recouvrent totalement 2 de ces 8 régions (Ile-de-France et Haute-Normandie), plus de 90% de 2 autres régions (Champagne et Picardie), 60% d'une région (Bourgogne) et moins de 30% des 3 autres régions (Centre, Basse Normandie et Lorraine). Les PRA sont de tailles très hétérogènes.

Dans la méthode de spatialisation des activités (LUCAS) exposée ci-dessus, la carte d'utilisation des sols est optimisée (« cross entropy ») sous la contrainte que la surface agrégée à l'échelle régionale soit égale à la surface estimée par ailleurs (par le RICA). Ainsi en re-agrègeant les surfaces des polygones (les USTI de la Figure 8) pondérées par la probabilité d'y trouver du blé (par exemple) vers la région entière, on obtient la valeur RICA de surfaces occupée par le blé dans cette région (donnée RICA 2002, année utilisée lors de la calibration GCE).

Les PRA étant des unités géographiques plus petites que la région, l'idée est d'utiliser ces données pour que la carte d'utilisation des sols soit plus proche de ce que l'on observe à l'échelle des PRA, ce qui ne peut qu'améliorer (au sens du rapprochement entre probabilités a priori et probabilités recalculées) la représentation de la couverture des sols, et donc par la suite la localisation des groupes type AROPAj. Il s'agit donc de bénéficier d'une couche d'information supplémentaire et de reprendre le processus de modélisation au stade « CE ». On substitue les valeurs PRA de la surface par type d'activité aux valeurs RICA (valeurs dénotées R_j dans le programme de minimisation de l'entropie croisée).

Les données PRA fournies sont datées de 2000. La Figure 10 montre la transformation de la carte d'occupation des sols lorsque l'on passe de l'utilisation des données RICA à celle des données PRA (application à la localisation des prairies permanentes en Picardie).

5.4. Passage aux mailles MODCOU

Différentes tâches imbriquées s'inscrivent donc dans ce programme de recherche. En matière de couplage des modèles AROPAj et MODCOU, tout comme ce qui concerne la relation STICS-AROPAJ, il s'agit de couplage « faible » ne nécessitant ni recodage informatique des modèles, ni transfert de code d'un modèle à l'autre. Il s'agit donc d'identifier les variables d'entrée et de sortie qui rendront ce couplage effectif. La réalisation de ce couplage suppose de systématiser et d'automatiser la méthode de désagrégation spatiale des résultats du modèle AROPAj à une échelle spatiale compatible avec le modèle MODCOU.

La transmission des « sorties » AROPAj vers MODCOU repose d'une part sur les probabilités de présence des groupes type AROPAj sur chacune des zones élémentaires AROPAj (USTI), et d'autre part sur la superposition des USTI et des mailles MODCOU. **Cette superposition est compliquée du fait qu'une maille MODCOU peut interférer avec des USTI appartenant à plus d'une Région et/ou à plus d'une classe altitudinale**³. Les mailles MODCOU sont des « mailles régulières » (des « carrés ») de 4 tailles différentes (des carrés de 1km, 2km, 4km ou 8km de côté), alors que les USTI sont des polygones composés de cellules élémentaires CLC de dimension 100m x 100m.

Considérons une « sortie » AROPAj notée V (évaluée en unité physique ou monétaire par unité de surface agricole utile). Le modèle AROPAj en fournit la valeur pour chacun des groupes type (V_k). Considérons l'ensemble des USTI « source » s interférant avec la maille MODCOU « cible » m et l'ensemble des groupes type k contribuant à l'activité agricole sur ces USTI. Les surfaces des USTI (Z_s) et des mailles MODCOU (T_m) sont connues. Avec la méthode de spatialisation des groupes type, on connaît la contribution de chaque groupe type

³ Dans les régions concernées par le PIREN, on ne rencontre que 2 classes d'altitudes (altitude>300m, 300<Altitude<600m, la classe Altitude>600 ne se rencontrent jamais sur le bassin de la Seine). Ces régions sont la Bourgogne, la Lorraine et la région Champagne-Ardenne.

k sur chaque USTI s , qui est la probabilité notée p_{sk} . La valeur de la sortie V par unité de surface de la maille MODCOU m s'obtient par la moyenne pondérée :

$$V_m = \sum_k \sum_s p_{sk} \frac{Z_s}{T_m} V_k$$

La Figure 11 illustre la relocalisation des groupes type d'une Région depuis les USTI vers les mailles MODCOU en l'absence d'effet de bord (zoom sur une zone « homogène » dans laquelle les USTI et les mailles MODCOU appartiennent à la même classe altitudinale et à la même Région). Il importe de remarquer que la localisation des groupes type sur les mailles MODCOU n'a pas d'utilité avérée au-delà de l'exemple de transfert de l'information d'un maillage à un autre.

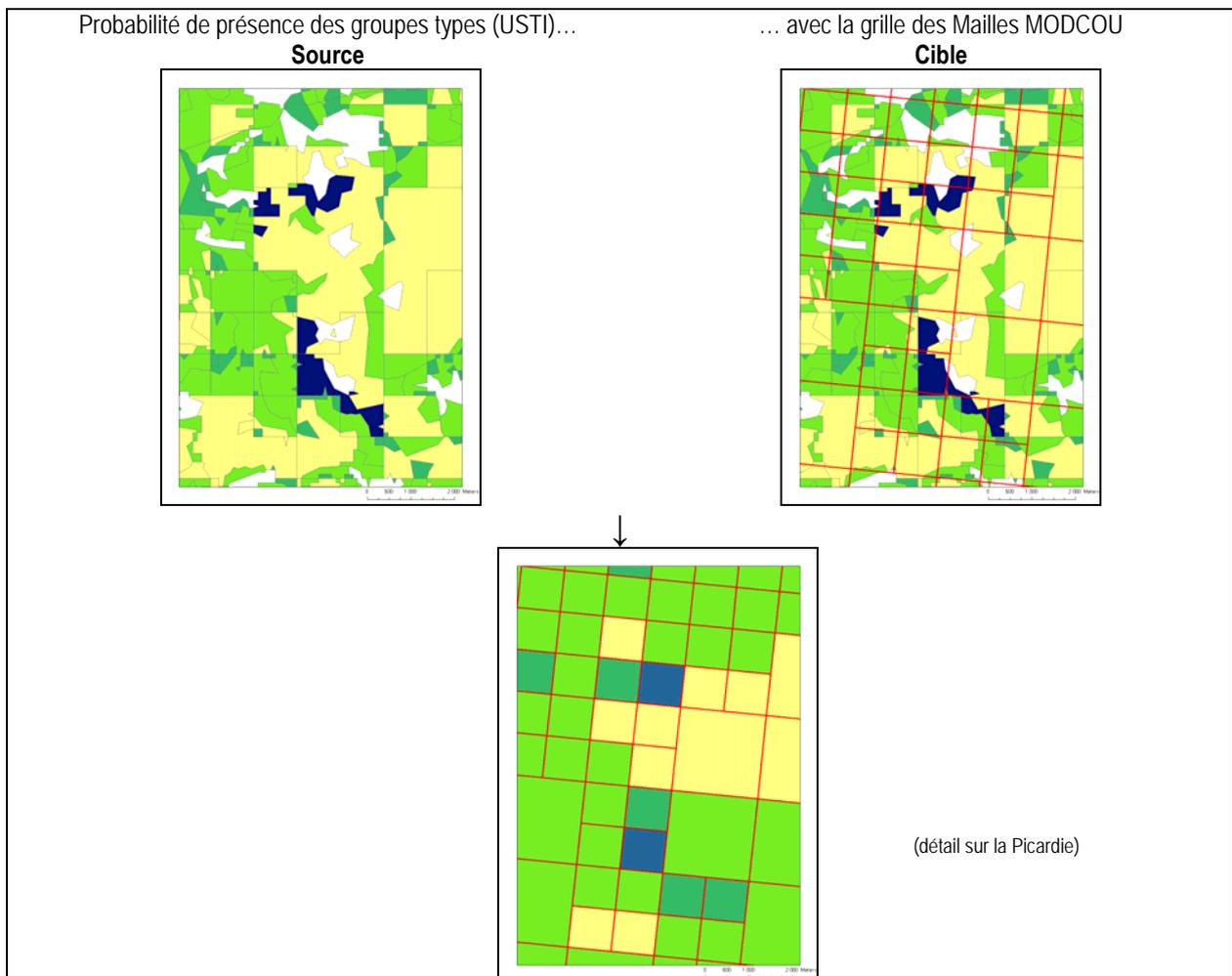


Figure 11 – Superposition des unités spatiales AROPAj (USTI) et d'une maille MODCOU.- exemple de la localisation probabiliste des groupes types sans « effet de bord ».

Dans le cas « hétérogène » d'une maille MODCOU débordant sur plus d'une Région, il faut tenir compte des contributions des USTI provenant des différentes régions au prorata des surfaces régionales sur la maille MODCOU. Prenons l'exemple de la représentation spatiale de l'émission de N_2O . La valeur de l'émission attribuée à une maille MODCOU à cheval sur 2 régions sera la moyenne pondérée des émissions de N_2O issues des USTI ayant une intersection avec la maille MODCOU. La pondération est définie par les parts de surface des USTI sur la maille MODCOU, quelque soit la Région d'appartenance des USTI.

Lorsqu'une maille MODCOU contient uniquement des zones déclarées *Non Agricole* (NAU), elle se voit attribuer la valeur *Non Agricole* (NAU) dans le seul cas où cette maille contient uniquement des zones déclarées comme tel sur la carte source des USTI. Dans le cas contraire (la maille MODCOU n'est pas couverte que d'USTI NAU), la partie Non Agricole est incluse dans l'interpolation par moyenne pondérée par sa surface dans la maille.

La Figure 12 montre la carte des probabilités du Groupe Type n°24 en Haute-Normandie avant et après interpolation sur la maille MODCOU.

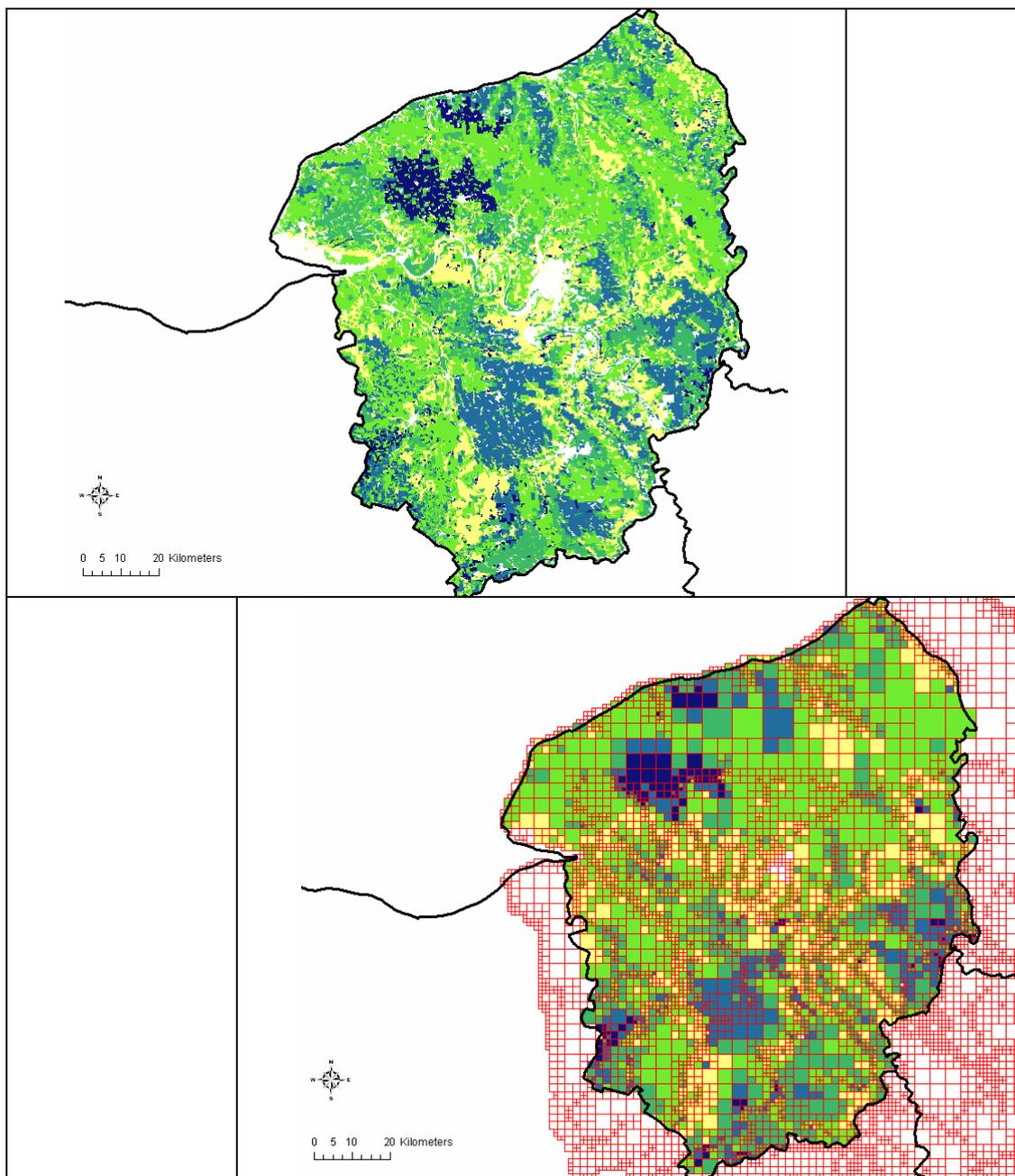


Figure 12 – Exemple d’interpolation des valeurs sur les polygones CLC vers les mailles MODCOU, région Haute-Normandie (RICA n.133). Les probabilités du groupe type 24 sont représentées sur cette figure.

6. Résultats pour l'estimation des émissions de polluants azotés d'origine agricole

6.1. Présentation des simulations test

Toute la section 6 est consacrée aux premiers tests de la méthode d'ensemble, intégrant :

- (i) les simulations AROPA_j dans la version V2 (RICA et typologie 2002 créées en 2005) améliorée pour le programme PIREN,
- (ii) (ii) avec la mise en œuvre des fonctions « dose-réponse » dans ce modèle,
- (iii) (iii) avec la mise en œuvre des modules de calcul des émissions de polluants azotés,
- (iv) (iv) la spatialisation des groupes type AROPA_j.

Le « scénario » retenu est le scénario de référence AROPA_j pour son calibrage (PAC et données 2002). Les résultats concernent les pertes d'azote sous forme de N₂O, de NO₃ (base racinaire), et de NH₃.

L'analyse des résultats « test » doit être faite en tenant compte des éléments suivants :

- Une simulation AROPA_j conduit généralement à une solution comportant des surfaces à valeur zéro parmi l'ensemble des cultures initialement retenues pour chacun des groupes type, même si la solution de « calibrage » devrait conduire en limiter le nombre. Considérons les 329 couples {groupe type *k*, culture *j*} résultant du couplage STICS-AROPAJ et pour lesquels sont disponibles les fonctions « dose-réponse ». L'optimisation AROPA_j ne conduit pas à 329 couples pour lesquels la surface est à valeur strictement positive. La solution optimale traitée dans l'exemple « test » montre que 317 couples parmi les 329 sont associés à des surfaces strictement positives.
- Il y a aussi des activités agricoles présentes (et à surface positive à l'optimum) pour lesquelles le couplage AROPA_j-STICS n'a pas été activé (prairies, fourrages, autres). Elles sont intégrées dans AROPA_j à rendement fixe, mais, sauf pour N₂O, polluant pour lequel on utilise le facteur d'émission IPCC par défaut, les émissions de pollution ne sont pas disponibles (et donc implicitement à valeur zéro).
- La spatialisation des variables AROPA_j lorsque celles-ci sont calculées de façon incomplète et hétérogène n'est donc ici produite qu'à titre d'illustration. Le Tableau 5 donne quelques indications sur les cultures présentes et les surfaces calculées à l'optimum, pour la totalité des 8 Régions et des 54 groupes type AROPA_j qui les composent (débordant donc au-delà du Bassin de la Seine pour 6 d'entre elles).

Région	Effectifs exploitants représentés (1000)	SAU (1000 ha)	Nombre de couples {k,j} AROPA _j -STICS avec surface positive en situation de référence	Part de SAU représentée par « AROPA _j -STICS »
121	3.8	489	22	88.7%
131	10.9	1406	45	74.4%
132	10.6	1162	44	79.3%
133	7.5	724	25	59.5%
134	19.8	2325	43	70.8%
135	16.5	1146	28	32.8%
136	13.1	1670	50	52.0%
151	8.1	1079	21	60.0%
Total	90.2	10000	278	63.7%

Tableau 5. Surfaces associées aux fonctions « dose-réponse » pour les 8 régions RICA affectant le Bassin de la Seine.

Dans les illustrations test que l'on propose, sont présentées les options alternatives renvoyant à des questions que l'on peut considérer comme « ouvertes » à ce stade de réflexion, par exemple :

- Evaluations « avec » ou « sans » constante dans les fonctions *apport d'azote – perte d'azote* ;
- Probabilités de localisation des groupes type AROPAj intégrant, ou n'intégrant pas, la couche d'information « PRA ».

On présentera également la transformation des résultats obtenue sur le maillage MODCOU.

6.2. Exemple de spatialisation de résultats de simulation

Une simulation AROPAj dans sa version V2 appliquée à la France (157 groupes type) requiert environ 20 à 40 minutes de temps calcul sur le serveur de calcul de l'INRA Grignon, selon l'état de charge du serveur. Une « sortie » AROPAj est un ensemble de « résultats » combinant « paramètres » et « variables d'optimisation » du modèle. Les sorties auxquelles on s'intéresse sont les pertes d'azote par hectare de SAU.

Le modèle AROPAj fournit les quantités de perte d'azote sous les formes respectives d'émission de N₂O, NH₃ et NO₃ pour chaque groupe type. Grâce au calcul des probabilités de présence des groupes types sur chaque USTI, on peut dresser une carte des émissions sur le bassin de la Seine. La Figure 13 illustre la distribution spatiale des émissions de N₂O estimées selon deux modalités de calcul, « avec » et « sans » la constante dans la relation entre apport et perte d'azote (voir la section 4.2). Cette figure utilise les probabilités de présence des groupes types calculées avec l'aide des PRA. L'exemple proposé montre que la spatialisation s'affranchit des frontières administratives malgré leur caractère fortement déterminant pour la typologie AROPAj. Il faut cependant remarquer que les émissions N₂O sont les mieux couvertes par le modèle, puisque leur évaluation intègre le fait que toutes les activités agricoles ne sont pas couvertes par des fonctions d'émission « dose-réponse ».

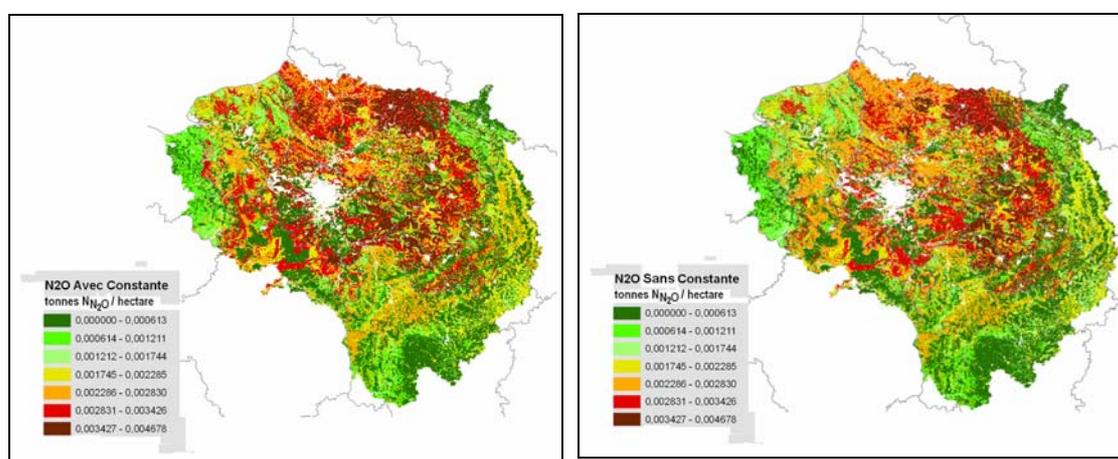


Figure 13. Spatialisation des flux d'émission de N₂O calculés par AROPAj (exprimés en tonnes de N sous forme N₂O par ha de SAU et par an) avec utilisation de l'information « PRA ».

6.3. Interpolation sur les mailles MODCOU

Compte tenu des dimensions moyennes respectives et de la dispersion des mailles sur l'espace géographique, le maillage MODCOU « dégrade » la qualité du « résultat » précédemment exposé avec le maillage des USTI. On présente sur la Figure 14 les émissions de N₂O interpolées sur les mailles MODCOU. Il est évident que cela n'a que valeur d'exemple, le maillage MODCOU étant surtout utilisé pour les pertes en nitrate.

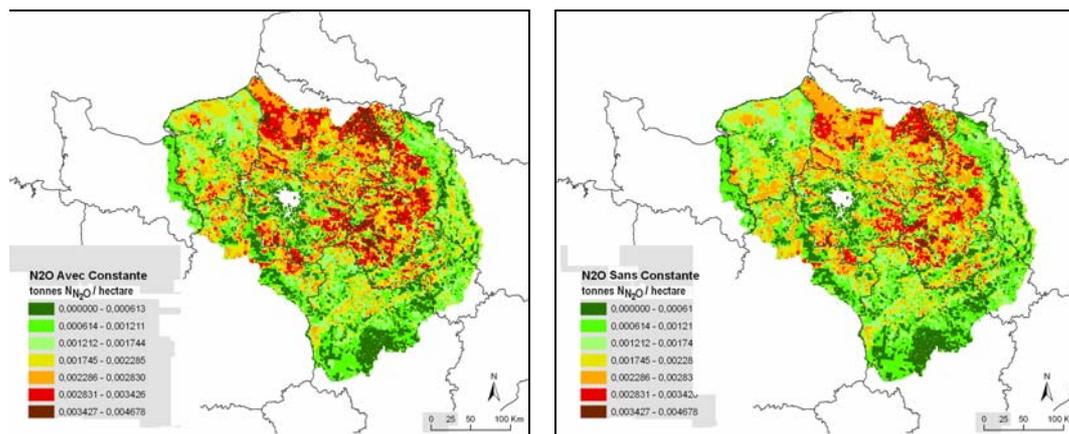


Figure 14. Transformation de la spatialisation par le maillage MODCOU : exemple des émissions de N₂O (« avec » et « sans » la prise en compte de la constante dans les fonctions affines d'émission de pollution).

6.4. Effet « PRA » sur la spatialisation des résultats

Les cartes de la Figure 15 présentent la spatialisation des pertes d'azote sous forme N₂O et NH₃ sans (à gauche) et avec (à droite) l'utilisation des PRA pour le calcul des probabilités de localisation des groupes types AROPA_j. Les émissions sont ici calculées sans intégrer les « constantes » dans les relations affines entre apport et pertes d'azote.

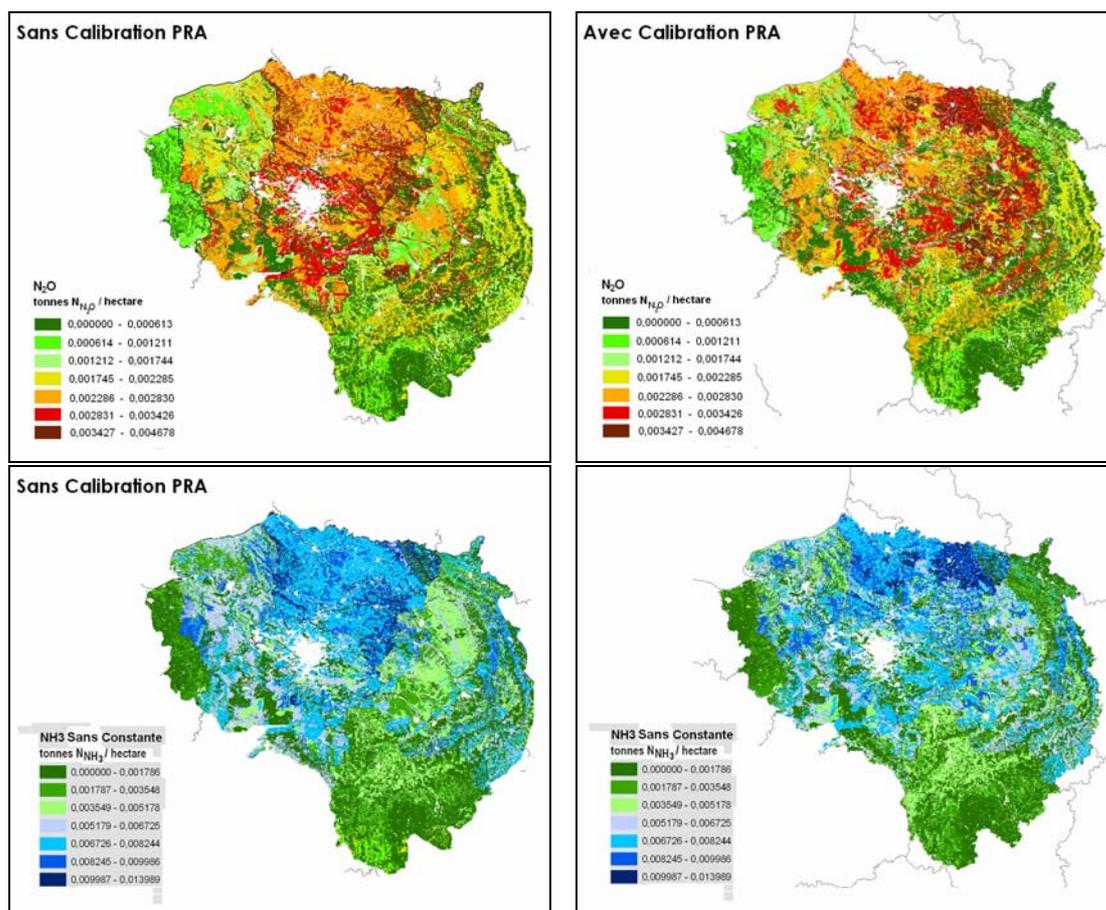


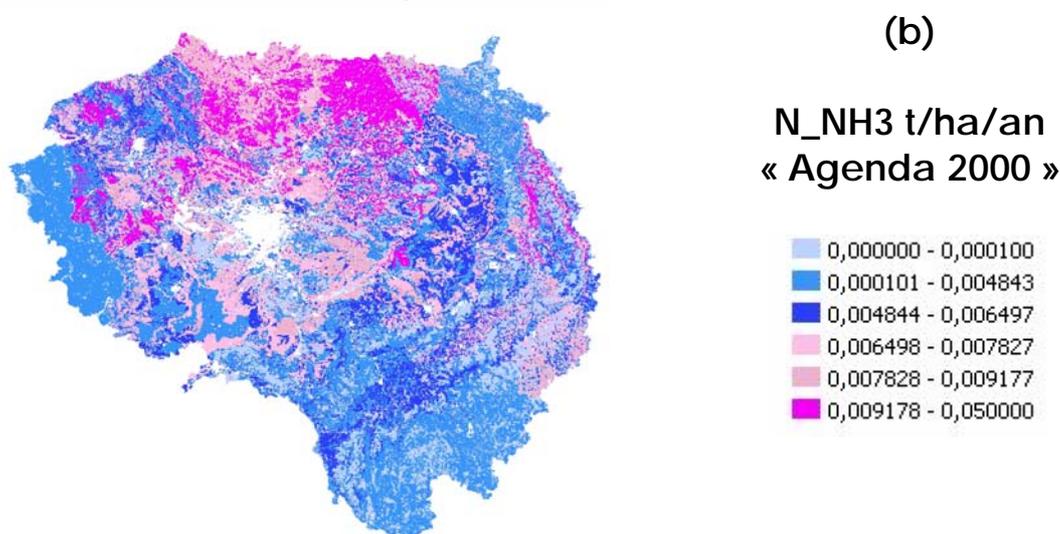
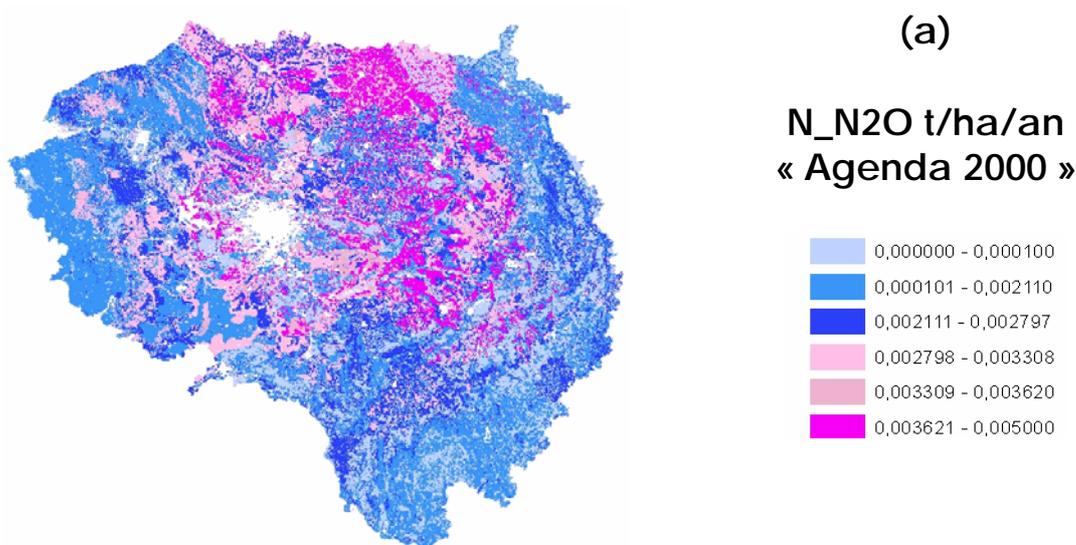
Figure 15. Impact de la prise en compte de l'information « PRA » sur la spatialisation des pertes d'azote sous les formes respectives N₂O et NH₃ (sans prise en compte de la constante dans les relations affines entre apport et pertes d'azote)..

6.5. Premiers résultats sur les pertes en nitrate

La vérification de certains paramètres du modèle STICS a été nécessaire, en particulier les paramètres relatifs à « l'imperméabilité » des sols inférés par le couplage STICS-AROPAJ lors du processus d'élaboration des fonctions de rendement. Les premiers résultats préalables à toute expérience de couplage AROPAJ-MODCOU sont illustrés par les Figure 16.

On montre en premier lieu (Figure 16 a, b et c) la distribution des pertes d'azote sous les 3 formes (N₂O, NH₃ et NO₃), calculées en tenant compte de la « constante » estimée pour les fonctions de perte (le terme B dans la relation $e(N) = A N + B$). Il importe de rappeler que les pertes ici calculées sont relatives aux seuls apports d'engrais minéral pour les seules cultures pour lesquelles sont disponibles les fonctions de réponse.

A ces illustrations correspondant au scénario AROPAJ de référence, à savoir l'année 2002 associée à la PAC en vigueur (« Agenda 2000 »), on ajoute pour le nitrate (Figure 16 d et e) les pertes sous forme de nitrate lorsque l'on modifie la politique agricole. Les scénarios retenus correspondent à l'accord de Luxembourg avant le relâchement de la contrainte sur le gel de terres, et à un scénario de découplage « complet » dans lequel la destination des terres n'est soumise à aucune contrainte et dans lequel tous les « recouplages » partiels (i.e. aides directes) autorisés par l'accord de Luxembourg sont supprimés. Il n'y a pas de variation de prix dans les scénarios retenus.



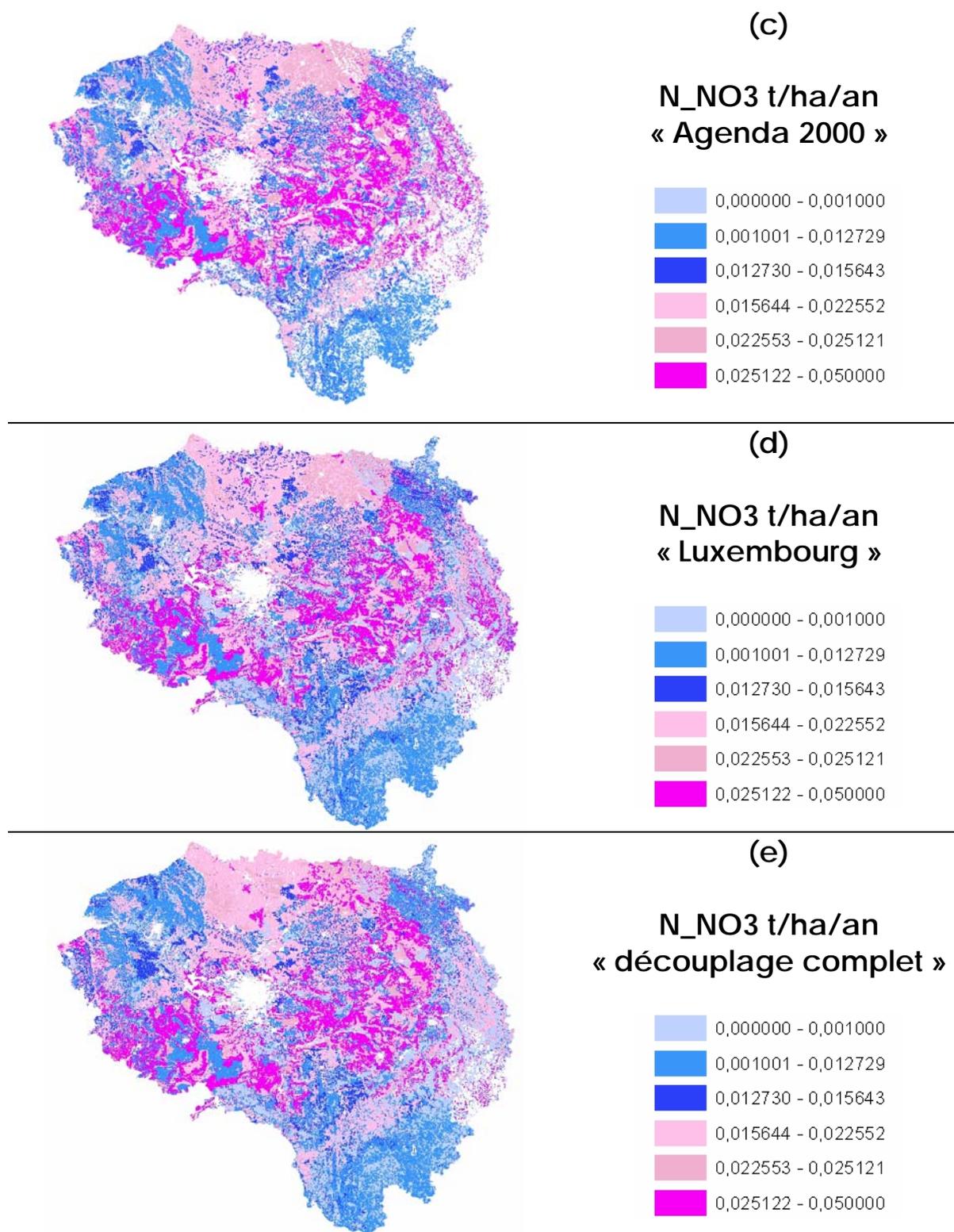


Figure 16. Pertes d'azote pour les postes « N2O » (a), « NH3 » (b) et « NO3 » (c) calculées dans différents contextes de politique agricole « Agenda 2000 » (c), « accord de Luxembourg » (d) et « découplage complet » (e), et exprimées en tonne d'azote par hectare de surface réelle et par an.

Ces illustrations ont conduit à de premières observations discutées lors du colloque PIREN-Seine 2009 quant à l'effet « complément » ou « substitut » des pollutions jointes, et quant à l'effet « allocation des terres » sur les pertes par hectare en dehors de tout effet « prix ».

7. Eléments de discussion technique

Ne sont ici discutés que les problèmes à caractère technique quant à la méthode d'ensemble ou quant aux aspects informatiques. La validation des premiers résultats n'est pas abordée.

7.1. Avantages et limites de la méthode et de son application

La pertinence des modèles utilisés dans leur champ propre d'application n'est pas ici discutée. On traite de l'intérêt que l'on peut trouver à associer le modèle économique AROPAj, le modèle agronomique STICS et le modèle hydrogéologique MODCOU, et de la pertinence à les associer de la façon proposée dans cette étude. On met ici en avant la dimension « opérationnelle » de la méthode.

7.1.1. Informatique et calcul

Les modèles sont en réalité associés en forçage, et non en couplage sous forme d'un modèle intégratif. Il n'est donc pas nécessaire ni utile de modifier (recoder) un modèle pour qu'il soit utilisable en relation avec les autres. La préservation de l'autonomie des modèles est de ce point de vue un avantage, puisque toute amélioration d'un modèle bénéficie aux autres sans qu'il soit besoin de modifier ces derniers. La difficulté devient celle de devoir traiter dans les formats adéquats les transferts d'information entre les modèles.

Le modèle AROPAj a été développé en environnement UNIX et cet environnement est requis pour utiliser l'ensemble des applications qui en permettent une utilisation simple par l'utilisateur. Cependant, les traitements numériques utilisant le compilateur FORTRAN ou le logiciel GAMS sont parfaitement transposables sous d'autres environnements (Windows). Le modèle STICS utilisé par l'UMR Economie Publique peut être compilé sous les deux environnements (avec néanmoins quelques difficultés selon le serveur de calcul). L'application ArTiX qui, dans sa dimension « base de données », associe des données AROPAj et des données STICS, a été transposée dans les deux supports (UNIX et Windows), avec quelques difficultés résiduelles tenant au logiciel de gestion de base de données utilisé. Une grande partie d'ArTiX fait appel au langage JAVA. ArTiX a été développé pour faciliter l'organisation automatique d'un très grand nombre de simulations STICS en vue de l'élaboration des fonctions de réponse (voir la section 4.1). Le modèle MODCOU (sous environnement LINUX) est en phase de recodage. Il peut être associé de façon intégrée à une version particulière de STICS.

L'utilisation d'un SIG pèse fortement sur la méthode générale de couplage, en particulier parce que les SIG standards n'en permettent pas l'usage en environnement informatique UNIX. On ne peut donc l'associer aisément de façon intégrée dans le processus de calcul. Par contre les phases de calcul pour la mise en œuvre des modèles d'économétrie spatiale sont réalisables sous les deux environnements UNIX et Windows (on utilise le logiciel « libre » R pour l'exploitation du modèle *MNL* et le logiciel GAMS pour l'exploitation du modèle *cross entropy*).

Sur le plan opérationnel, il convient de distinguer de ce qui relève du couplage proprement dit, et de ce qui relève de l'élaboration et du calibrage des modèles amont. Entre ces deux aspects, il y a les développements particuliers réalisés pour le compte du PIREN-Seine.

Les moyens les plus importants sont mis en œuvre pour l'élaboration et le calibrage des modèles, réalisés pour tout un ensemble de programmes qui dépassent le PIREN-Seine. Il est vrai également que le PIREN-Seine conduit à des développements utiles à d'autres programmes. Bien que cela dépende fortement de la qualité du serveur de calcul (le serveur utilisé à Grignon semble dépassé largement quant à sa vitesse de calcul), on peut donner des ordres de grandeurs pour ce qui concerne AROPAj : la typologie (automatique) et la pré-estimation (automatique) des paramètres à partir du RICA requièrent 3 jours de calcul pour

l'ensemble de l'UE ; le calibrage (automatique) d'une version à l'échelle de l'UE requiert environ 1 semaine de temps calcul pour 100 groupes type ; une simulation demande environ 20 minutes pour les 157 groupes type français. La mise à disposition de l'ensemble des plus de 700 fonctions dose-réponse obtenues pour la France à partir de STICS-ArTiX a demandé plusieurs semaines. Le modèle *cross entropy* requiert environ 1 jour de temps calcul par Région (le facteur limitant est la dimension géographique plutôt que le nombre de groupes type). Ces temps n'intègrent évidemment pas la phase amont de la conception proprement dite des modèles.

Nous ne disposons pas d'une chaîne complète de traitement, intégrant les simulations AROPAj et les simulations MODCOU. Il est donc difficile d'évaluer les temps calcul nécessaires à la réalisation d'un scénario complet. Le premier facteur limitant pourrait être le temps requis pour les simulations MODCOU si celles-ci doivent intégrer des simulations STICS. Si le « signal » provenant d'AROPAJ est considéré comme satisfaisant, on peut envisager d'utiliser MODCOU sans le module STICS qui peut lui être associé. Le second facteur limitant vient de l'articulation entre le modèle économique et le modèle hydrogéologique dès lors qu'intervient l'interpolation spatiale des résultats AROPAj. En effet la phase de calcul par SIG suppose d'extraire les « sorties » AROPAj en amont pour les spatialiser dans un environnement informatique différent encore peu propice au traitement automatisé.

Si l'on s'oriente vers le calcul des pertes en nitrate via la voie « STICS-AROPAJ », l'avantage de préserver le lien « géo-référencé » entre choix et activité économique d'une part et pollution d'autre part, deux problèmes sont à traiter. Le premier problème est celui de la distribution temporelle, sur un pas de temps MODCOU (quotidien), des pertes en nitrate calculées annuellement par AROPAj. L'idée retenue est la suivante : on « cale » (pour chaque maille MODCOU) la distribution AROPAj obtenue pour l'année de calibrage AROPAj sur la distribution calculée par le couplage STICS-MODCOU pour la même année. Pour les autres années éventuellement requises par un scénario « dynamique » (voir le début de la section 3), on extrapole les profils infra-annuels calculés par STICS-MODCOU aux résultats AROPAj. Cela suppose que les fonctions « dose-réponse » aient été ajustées pour ces années climatiques (comme cela fut réalisé par Galiko-Debove, 2007).

7.1.2. Résolution spatiale et frontières

Le caractère déterminant des Régions RICA dans le modèle AROPAj, avec l'effet d'homogénéisation de l'agriculture régionale, transparait évidemment dans la spatialisation. Il est difficile en amont, au niveau des données RICA, de jouer avec cet effet, puisque le droit d'accès aux données interdit toute localisation des exploitations agricoles de l'échantillon d'enquête. Au sein d'une Région, la seule information exploitable à des fins de localisation est « l'altitude » (en réalité résumée par les 2 seuils de 300 et 600m), en général peu discriminante dans la plupart des Régions. Les données économiques ne permettent donc aucune localisation possible a priori des groupes type.

Le véritable lien entre les agents économiques et l'espace géographique est créé à partir des activités agricoles qu'il est possible de localiser (via les bases de données CLC et LUCAS). Un lien existe entre l'activité agricole et le type de sol, à travers le couplage entre STICS et AROPAj, mais il paraît encore ténu et difficilement exploitable à des fins de localisation. Il est de plus compliqué par la difficulté de faire coïncider à un haut niveau de résolution la succession de cultures sur un même sol telle que les études agronomiques la proposent, et les parts de surface admissibles affectées aux cultures par AROPAj qui reposent sur un calcul statistique revu par la phase de calibrage de ce modèle.

A l'observation des cartes présentées comme des illustrations de l'approche proposée, il apparaît clairement que l'information sur la localisation des activités agricoles affinée par les

PRA a un impact significatif sur la qualité de la spatialisation. Les artefacts que constituent les frontières administratives sont sensiblement moins présents lorsque la spatialisation des groupes types (et conséquemment des « sorties » AROPAj) est calibrée en utilisant les PRA comme source additionnelle d'information.

7.1.3. Validation des fonctions de réponse

L'observation du faisceau des courbes de rendement en blé (présent chez un grand nombre de groupes type) en réponse aux apports d'azote (voir Figures 5) fait apparaître la très grande variabilité de la « réponse » au sein d'un petit groupe de Régions voisines. Cette variabilité existe aussi au sein même des Régions. La « courbure » des fonctions de rendement joue un grand rôle dans le calcul du rendement optimal (voir Figure 1), même lorsque le « potentiel agronomique révélé par la valeur asymptotique différencie peu les fonctions. Cela se traduit par des écarts significatifs dans les apports d'azote par unité de surface, et donc par des pertes d'azote sensiblement différentes. A ce stade c'est le choix de la forme fonctionnelle de la fonction de rendement qui est en jeu. On notera cependant que la fonction choisie est une de celles que les agronomes retiennent, que cette fonction satisfait aux conditions requises par l'économie (monotone croissante et concave) et que l'ajustement avec les « sorties » STICS apparaît satisfaisant. Par ailleurs, le caractère linéaire des fonctions de pollution est également un « standard » (du moins, si l'on suit les travaux de l'IPCC, pour la fonction « N2O » qui est celle pour lequel ce caractère semble le moins évident au vue des « sorties » STICS).

Les fonctions « dose-réponse » contribuent donc à différencier fortement d'une région à l'autre la distribution de certaines sorties AROPAj, comme le montrent les représentations cartographiques que l'on peut en donner. La « sortie » NO3 est celle qui cumule les handicaps, alors qu'elle est l'objet du couplage entre MODCOU et AROPAj :

- Beaucoup des paramètres des fonctions affines de pollution NO3 sont estimés à zéro ;
- Les surfaces agricoles pour lesquelles on dispose de fonctions de pollution sont hétérogènes, en part de surface, d'une région à l'autre ;
- On ne dispose pas de valeur par défaut à proposer pour NO3, comme il en existe pour N2O (valeur IPCC).

La première difficulté soulevée ici renvoie à la validation de STICS (par l'expérimentation ou par le choix du formalisme), ou à la façon avec laquelle ce modèle a été utilisé (en particulier au niveau de la succession des cultures et de l'évolution de l'azote stocké dans le sol). Les deux autres difficultés pourraient être levées à partir d'une méta-analyse analogue à ce qui a été proposée pour l'émission de N2O (cf les travaux de Bouwman⁴). Dans les deux cas, la solution ne relève pas à proprement parler de l'analyse économique. Il reste que l'on peut tirer profit du travail de couplage réalisé entre les modèles STICS et MODCOU, et qu'il est possible de « recalibrer » les fonctions de réponse en combinant les observations « moyennes » habituellement faites (sur les pertes d'azote) et les « sorties » STICS dont nous disposons.

7.2. Questions ouvertes

La méthode générale permettant d'estimer l'impact d'un changement de l'environnement économique sur la distribution des pertes d'azote en nitrate sur l'espace géographique du Bassin de la Seine est techniquement opérationnelle. La transmission du « signal » obtenu vers MODCOU suppose que soient abordés ou revus les points « amont » ou « aval » suivants. Les points « amont » renvoient aux problèmes soulevés en section 4.2. Les points « aval » concernent les modalités de transfert vers MODCOU.

⁴ Bouwman, A., (1996), Direct emission of nitrous oxide from agricultural soils, *Nutrient Cycling in Agrosystems*, 46, pp53-70.

➤ Amont

1. *Intégration de la « constante B » dans le calcul effectif des pertes Y sous la forme $Y = AX + B$.* Ne pas l'intégrer peut avoir un sens pour N₂O si l'on considère que ce qui importe ce sont les émissions additionnelles d'origine anthropique et que la constante est interprétée comme la part naturelle des émissions. Ce dernier point est lui-même critiquable dans la mesure où les émissions dépendent du type de culture, part naturelle comprise, et que le choix des cultures relève bien d'une décision humaine. Il est difficile d'ignorer la « constante » en ce qui concerne les nitrates, puisque naturels ou pas, les excès d'azote en nitrate au niveau des sols vont modifier la concentration en nitrate des aquifères, qui est le critère de qualité généralement retenu.
2. *Fonctions de réponse et climat.* Toutes les simulations STICS le montrent, la réponse en production ou pollution à l'apport d'azote dépend fortement du climat. Les simulations STICS que nous avons utilisées sont en rapport avec les données économiques, et l'année de base des observations est donc 2002. Le fait que l'année 2002 n'apparaisse pas du point de vue climatique comme une année accidentelle à l'image de l'année 2003 ne joue pas contre le fait que l'on puisse la considérer comme « moyenne ». Au-delà des problèmes informatique et d'accès aux données que suppose la généralisation des simulations STICS pour tout un jeu de « climats » (probabilisables ?), émerge la difficulté de « résumer » la variabilité par un nombre limité de paramètres qui permette l'exploitation de ces simulations par le modèle économique.
3. *Choix des agents économiques dans l'incertain.* L'hypothèse de comportement des agents économiques dans le modèle AROPAj signifie que les agents fondent leurs décisions comme s'ils connaissaient à la fois le climat et l'environnement économique pour la campagne de production (sur 1 an). Les choix dans l'incertain ont donné lieu à une abondante littérature économique théorique et à des développements en matière de recherche opérationnelle. Testées au début du développement du modèle AROPAj, la formalisation des choix dans l'incertain et la mise en œuvre informatique ont été écartées dans la mesure où le déplacement des difficultés (l'estimation des paramètres associés à l'aversion pour le risque) et le « coût » des calculs ne nous paraissaient pas compenser la qualité espérée des résultats.
4. *Ré-estimation des paramètres des fonctions d'émission de pollution.* Cette possibilité est assez facile à mettre en œuvre. Elle demande un investissement dans le traitement des « sorties » STICS cumulées, alors que les procédures de traitement automatique existant (ArTiX) reposent sur l'itération des étapes de calcul indexées sur le groupe type AROPAj et la culture.
5. *Re-calibrage des paramètres des fonctions d'émission de pollution.* Sans rapport avec le point précédent, l'idée est de recalculer l'ensemble des paramètres de sorte que la valeur moyenne des émissions par unité d'azote apportée corresponde à une valeur donnée (cette procédure semble avoir été proposée en matière de calcul des émissions de N₂O). La difficulté reste néanmoins de faire en sorte que les bilans élémentaires d'azote (bilans massiques, sur la base des sorties d'azote sous forme de protéine et de « pertes » identifiées) soient respectés à l'échelle de la parcelle.
6. *Reprise de la notion même de perte d'azote.* Il a été suggéré (Céline Schott, INRA-Mirecourt) que la perte d'azote pouvait être comprise comme l'écart entre la dose « optimale » d'azote apporté et la quantité effectivement apportée. La dose optimale est alors définie comme la dose minimale requise pour obtenir le potentiel de rendement espéré avec des conditions agronomiques (pédo-climatiques ?) données.

L'excès d'azote peut venir d'une mauvaise anticipation des conditions climatiques et d'un excès de précaution face aux aléas. Il reste à discuter de ce qui relève de la décision de l'exploitant agricole et de la façon avec laquelle on traduit cette prise de décision dans le modèle économique (voir le point 3 ci-dessus), et de la répartition des pertes entre les différents postes (N₂O, NO₃, ...).

➤ Aval

7. *Distribution du signal « AROPAj » spatialisé sur l'année au pas de temps MODCOU.* AROPAj fournit une « sortie » annualisée, qu'il convient de redistribuer sur l'année en tentant de rapprocher cette distribution de celle qui prévaudrait dans le déroulement du développement d'une culture sur un cycle de production.
8. *Re-calibrer le signal AROPAj sur la base des estimations issues du couplage STICS-MODCOU.* L'idée est de comparer les pertes en nitrate obtenues par la méthode développée dans ce papier, avec les pertes calculées par le couplage STICS-MODCOU, pour chacune des mailles MODCOU et pour une année dite de « référence » (i.e. 2002). Il s'agirait alors de « redresser » en conséquence les sorties AROPA à l'échelle de la maille MODCOU.
9. *Format et organisation des transferts entre AROPAj et MODCOU.* L'élaboration des fichiers informatiques pour un test de transfert depuis la « sortie » AROPAj vers l'« entrée » MODCOU ne devrait pas présenter de difficulté. Mais la gestion d'une succession de « transferts » de ce type demandera un investissement à la mesure du type de forçage envisagé entre les modèles, du nombre d'itérations ou de répétitions requises pour la réalisation d'un scénario complet (i.e. 30 transferts pour 30 années, éventuellement démultipliés par les options permises à chaque itération : par exemple en matière de « taxe nitrate »).

7.3. Poursuite du travail engagé

Les relations entre l'activité économique agricole et ses effets sur le milieu physique sont analysées à travers la modélisation, et plus précisément en mobilisant trois modèles :

- un modèle d'offre agricole : AROPAj ;
- un modèle biophysique : le modèle de culture STICS ;
- un modèle hydro-géologique : le modèle MODCOU.

Le modèle d'offre agricole AROPAj sera utilisé pour transformer des scénarios économiques⁵ en scénarios de changement d'usage des sols agricoles et de changement d'émission de pollution. Il s'agit d'intégrer ou de reprendre les modules de calcul des impacts environnementaux et les nouveaux instruments de régulation dans les scénarios que le modèle est capable de traiter. Pour l'UMR d'Economie Publique, l'analyse de la régulation économique portant sur des pollutions jointes à l'origine d'externalités différentes (pollution globale ou locale sur des horizons de temps différents selon qu'il s'agisse de N₂O, NH₃, NO₃) sera l'objet de la thèse financée par le programme R2DS de la Région Ile-de-France PIREN (projet « Eco-Seine »). A cet effet, Cyril Bourgeois est recruté en qualité d'allocataire en thèse à compter du 1^{er} décembre 2008. Le volet économique des pollutions des aquifères du Bassin de la Seine par le nitrate sera ainsi abordé en contrepoint au volet physique développé dans le cadre de la thèse d'Elodie Philippe.

⁵ On entend par « scénario économique » un scénario relevant de la Politique Agricole Commune, d'une politique de promotion des cultures énergétiques, ou de politiques environnementales. On peut si nécessaire élargir la notion de scénario à toute évolution de l'environnement physique qui s'impose au secteur agricole, comme cela peut être le cas avec l'évolution du climat. En matière de changement climatique, il convient donc de distinguer ce qui relève d'une part d'une politique de régulation des émissions de gaz à effet de serre (qui peut s'appliquer à tout secteur « contrôlé » par une autorité de régulation), et d'autre part ce qui relève de l'adaptation du système productif (éventuellement « encadrée » par une politique publique).

Bibliographie

- Baranger E., Jayet P.A., Roger-Estrade J., Richard G. (2008), **Coupling a biophysical and a micro-economic model to assess environmental and economic consequences of agricultural soil degradation due to compaction**, 29p.
- Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M. H., Ruget, F., Nicoullaud, B., Gate, P., Tayot, X., Plenet, D., Cellier, P., Machet, J. M., Meynard, J. M., and Delecolle, R. (1998). **STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn.** *Agronomie* 18, 311-346.
- Chakir R., (2009), **Spatial Disaggregation of Agricultural Production Data: An Econometric Approach using Minimum Cross Entropy**, à paraître dans *Land Economics*,
- De Cara S., Houzé M., Jayet P.A., (2005), **Green house gas emissions from agriculture in the EU: a spatial assessment of sources and abatement costs**, *Environmental And Resource Economics*, vol. 32, n° 4 . pp 551–583.
- Durandeu S., Gabrielle B., Godard S., Jayet P.A., Le Bas C., (2008), **Coupling biophysical and micro-economic models to assess nitrous oxide emissions from cropland and the effects of mitigation measures on greenhouse emissions**, 54p., en révision
- Godard C., (2005), **Modélisation de la réponse à l'azote du rendement des grandes cultures et intégration dans un modèle économique d'offre agricole à l'échelle européenne - Application à l'évaluation des impacts du changement climatique**, 13 décembre 2005, ABIES-AgroParisTech
- Godard C., Roger-Estrade J., Jayet P.A., Brisson N., Le Bas C., (2008), **Use of the available information at a European level to construct crop nitrogen response curves for the regions of the EU**, *Agricultural Systems*, vol. 97, n° 1–2, pp. 68–82. Avril 2008
- Galko-Debove E., (2007), **Modélisation de l'offre agricole européenne face à de nouveaux enjeux : réformes politiques, effet de serre et changement climatique**, 24 octobre 2007, ABIES-Engref.
- Jayet P.A., (2006), **Report on results concerning models linking farm, markets and the environment**, GENEDEC Deliverable 4, décembre 2006, 305p.
http://www.grignon.inra.fr/economie-publique/genedec/publi/deliv/WP3_D4.pdf
- Ledoux, E., Girard, G. & de Marsily, G., 1989: **Spatially distributed modeling : conceptual approach, coupling surface water and groundwater**. In : *Unsaturated flow in hydrologic modeling, Theory and practice*, Kluwer Academic Publishers, pp. 435-454.
- Schott C., C. Mignolet, M. Benoit, 2009 **Agriculture du bassin de la Seine, Découvrir l'agriculture du bassin de la Seine pour comprendre les enjeux de la gestion de l'eau**, Rapport # 5 de l'Agence de l'Eau Seine Normandie et du PIREN Seine, 79p
- Viennot P., E. Ledoux, J.M. Monget, C. Schott, C. Garnier, N. Beaudoin 2009 **La pollution du bassin de la Seine par les nitrates: comprendre l'origine et la migration des nitrates dans l'écosystème pour mieux protéger les aquifères** Rapport # 3 de l'Agence de l'Eau Seine Normandie et du PIREN Seine, 43p

Glossaire

Définition et explication des termes et acronymes qui suivent sont fournis en rapport avec l'usage qui en est fait directement ou indirectement dans ce document. Ces éléments sont fournis dans la mesure où les travaux qui sous-tendent l'étude relèvent de plusieurs disciplines scientifiques. On donne aussi la « signification » de quelques acronymes.

Agent économique

Un agent économique est défini par les choix économiques qu'il peut prendre et par les conditions qui s'imposent à lui. Les choix sont résumés par ses variables de « commande » (pour un agriculteur : niveau d'apport d'engrais, allocation des terres dont il a la gestion). On privilégie dans cette étude ses choix de « court terme »

AROPAj

Agriculture, Régulation de l'Offre et Politique Agricole : modèle économique d'offre agricole, développé à l'UMR d'Economie Publique INRA-AgroParisTech à Grignon. C'est un modèle de

programmation mathématique fondée sur l'optimisation des choix d'agents économiques représentant le secteur agricole, sous contraintes techniques (formant les « ensembles de production ») et réagissant à l'environnement économique (la PAC, les prix dans l'hypothèse d'agents « *price takers* »). L'agent élémentaire est le « groupe type » de producteurs. Un programme mathématique d'optimisation est associé à chacun des groupes type. Un tel programme se résume de façon formelle par l'expression : $\max_x F(x;a) : x \in X(a)$. AROPAj relève essentiellement de la programmation linéaire : la fonction d'objectif F et les contraintes définissant l'ensemble X sont des fonctions affines de x . Ce modèle est implanté sur un serveur de calcul en environnement UNIX. Sur le plan informatique, la programmation linéaire est mise en œuvre en deux temps : (1) construction de la matrice associée à F et X (en « ligne » l'objectif F les contraintes X ; en « colonne » les variables x) pour le jeu de valeurs des paramètres a ; (2) résolution du programme d'optimisation. On dispose pour cela de « générateurs de programmes et de matrices » et de « solveurs ». Pour la partie « réalisation des simulations » (après les étapes d'élaboration de la typologie et d'estimation des paramètres à partir du RICA), les dernières versions d'AROPAJ combinent des modules de gestion utilisant le script UNIX, des routines FORTRAN et le logiciel GAMS.

Lien : <http://www.grignon.inra.fr/economie-publique/MIRAJE/model/detail.htm>

Cellule élémentaire, maille élémentaire, pixel

Par abus de langage (pixel) et par commodité, on utilisera l'un de ces termes pour désigner la plus petite unité spatiale utilisée pour le croisement des données manipulées. Il s'agit en réalité de la « maille » CLC carrée de 100m de côté, et géo-référencée.

CLC

La base de données géographiques CORINE Land Cover (CLC) est produite dans le cadre du programme européen CORINE, de coordination de l'information sur l'environnement. La continuité de ce programme et la diffusion des données sont pilotées par l'Agence européenne pour l'environnement. En France, l'IFEN est chargé d'en assurer la production, la maintenance et la diffusion. Cet inventaire biophysique de l'occupation des terres fournit une information géographique de référence pour 29 Etats européens et pour les bandes côtières du Maroc et de la Tunisie. CLC décrit la couverture des terres sous forme d'une base géographique. À partir d'images satellitaires, le territoire est analysé comme une somme de polygones distinguant forêts, tissu urbain, milieux naturels, etc. Une première version de la base, dite CLC 1990, a été réalisée à partir d'images acquises entre 1987 et 1994. La base de données CORINE Land Cover 2000, dite CLC 2000, a été elle réalisée à partir d'images satellitaires de l'année 2000. Le programme CORINE Land Cover repose sur une nomenclature standard hiérarchisée à 3 niveaux et 44 postes répartis selon 5 grands types d'occupation du territoire : artificiel, agricole, forêts et milieux semi-naturels, zones humides et surfaces en eau.

Econométrie spatiale

La discipline se réfère la mise en relation de variables que l'on cherche à expliquer avec des variables explicatives. L'indexation est essentiellement ici spatiale (l'économétrie mobilise le plus souvent des variables indexées sur le temps). Les variables expliquées sont à mettre en relation avec la modélisation et le calcul économique, qu'il s'agisse de variables exprimées en valeur (revenu, prix, productions, consommations) ou en unité physique (surfaces, productions, pollutions). Il s'agira principalement de « surface » et d'activité liée à la surface. Les variables explicatives sont ici les caractéristiques physiques disponibles dans les bases de données géo-référencées européennes. Les modèles d'économétrie spatiale utilisés dans ce rapport sont le modèle *logit multinomial* et le modèle de *minimisation de l'entropie croisée*.

Fonction dose-réponse

Une fonction dose-réponse (ou fonction de réponse) met en relation un ou plusieurs input(s) et un output. La fonction peut être multidimensionnelle (plusieurs outputs). Dans le présent document, en application à la production agricole, l'input est l'apport d'azote et les outputs sont le rendement et différents polluants azotés. Le choix d'une forme fonctionnelle confère une forme analytique à la fonction, qui peut être obtenue à partir d'observations ou à partir de simulations fournies par un modèle biophysique.

Fonction de pédotransfert

Ces fonctions permettent d'associer les caractéristiques constitutives des sols et leurs propriétés

structurales. L'utilisation de ces fonctions est un préalable pour passer des données « sol » disponibles dans la Base Européenne des Sols aux paramètres nécessaires au fonctionnement du modèle *STICS* (étape réalisée par l'Unité Infosol de l'Inra Orléans) en association avec *AROPAj* pour élaborer les fonctions « dose-réponse ».

Géostatistique

Abordée ici dans le sens étroit d'une technique de distribution spatiale d'informations, visant à associer des données de sources multiples, à produire des informations géo-référencées, à cartographier ces informations, à l'aide d'un système d'information géographique. L'association des données peut utiliser des modèles relevant de l'économétrie spatiale.

Groupe type (AROPAj)

Un groupe type est un agent élémentaire auquel on prête la capacité de décider de ses choix de production. Il est supposé « rationnel » (et plus précisément fondant son choix sur un programme de maximisation du revenu tiré de son activité). Les groupes type sont déterminés par une technique de classification automatique fondée sur quelques critères disponibles dans le RICA. Les paramètres caractérisant chacun des groupes type (matérialisant les contraintes et la fonction d'objectif du programme d'optimisation de chaque groupe type) reposent sur des techniques d'estimation et de calibrage développées spécifiquement pour *AROPAj*.

HSMU

Homogenous Spatial Modelling Unit : équivalent de l'USTI utilisé dans d'autres travaux européens de spatialisation.

LUCAS

Le projet Européen Land Use/Cover Area frame statistical Survey (LUCAS), expérimenté en 2001, 2003 et 2005, a pour principal objectif d'obtenir des données harmonisées à l'échelle européenne sur l'occupation et l'utilisation du sol et leurs changements, avec une précision de 2 % ou plus pour les catégories majeures. Il repose sur un sondage à 2 phases (une enquête de terrain au printemps et un entretien avec les fermiers/exploitants agricoles en automne). L'échantillonnage du territoire pour ce sondage est effectué comme suit : un maillage carré de 18 km de côté et d'abord constitué puis à chaque nœud une grappe de 10 points (2 lignes de 5 points). Ceci représente environ 100 000 points visités à chaque enquête, répartis sur le territoire de l'Union Européenne à 15 (UE-15). La nomenclature des données LUCAS est très détaillée.

MARS

Monitoring of Agriculture by Remote Sensing : l'objectif du système agrométéorologique de suivi des cultures et de prévision des rendements, développé par le Centre Commun de Recherche de la Commission Européenne (CCR, ou JRC) dans le cadre du projet MARS, est de fournir les éléments de compréhension de l'impact des événements climatiques de la campagne sur les récoltes et d'anticiper des prévisions de rendement des principales cultures. Dans ce cadre l'une des missions du projet est l'achat de données météorologiques, traitement, calculs, interpolations et stockage des données météorologiques dérivées. Cela a permis la création d'une base de données météorologiques européenne, dénommée base de donnée MARS.

MNT

Modèle Numérique de Terrain (Digital Elevation Model DEM en anglais) ; représentation de la topographie (altimétrie et/ou bathymétrie) d'une zone terrestre sous une forme adaptée à son utilisation par un ordinateur numérique

MODCOU

Modèle hydro-géologique d'écoulement des eaux, développé au Centre Géosciences de l'Ecole des Mines de Paris.

Paramètres (modèle)

Les paramètres d'un modèle caractérisent l'environnement tel qu'il est représenté dans le modèle, de la même façon que les paramètres a d'une fonction analytique $f(x ; a)$ se distinguent des variables x de la fonction. Dans *AROPAj*, on distingue d'une part les paramètres caractérisant les groupes type et les paramètres généraux (i.e. la PAC et sa mise en œuvre dans le modèle), et d'autre part les variables d'optimisation. La notion de paramètre recouvre des éléments de nature différente, même s'ils jouent

un même rôle sur le plan mathématique. Certains paramètres sont des caractéristiques des groupes types (la SAU, le nombre d'exploitants représentés, les quotas, les charges variables, les rendements, ...); d'autres résultent des politiques agro-environnementale (PAC ; politiques de régulation mises en œuvre par exemple par l'instauration de taxes).

PRA

Petite Région Agricole : Unité réputée homogène sur le plan des conditions pédo-climatiques et des activités agricoles. Les PRA sont des zones contiguës formant des partitions départementales (et a fortiori régionales). Les **régions agricoles** et **petites régions agricoles** ont été définies (en 1946 et 1949, par l'INSEE) pour mettre en évidence des zones agricoles homogènes. La Région Agricole (RA) couvre un nombre entier de communes formant une zone d'agriculture homogène. **La Petite Région Agricole** est constituée par le croisement du département et de la RA. Par définition, dans une Petite Région Agricole on retrouve une même vocation agricole dominante.

Programme linéaire (optimisation)

De façon schématique un programme linéaire (PL) d'optimisation dans sa forme réduite s'écrit : $\max_x p(a)x : Lx \leq b(a) ; x \geq 0$. Dans l'acception économique, on y distingue le vecteur des variables primales x , le vecteur des « prix » p , la matrice des contraintes L , le vecteur des « ressources » b et le vecteur des paramètres a . Les solveurs standards donnent la solution complète (valeurs à l'optimum de toutes les variables primales et duales).

Rendement

La notion est utilisée dans le sens étroit de l'acception agro-économique. Pour la production végétale rapportée à la terre, elle exprime la quantité de produit commercialisable (i.e. le grain) obtenue par unité de surface utilisée. Dans AROPAj, les rendements agricoles sont indexés sur la culture et le groupe type. Ce sont a priori des paramètres du modèle. Lorsqu'on utilise les fonctions de réponse associant rendement et apport en azote, on remplace un « point » (apport, rendement) par une « fonction ». Les charges en azote et les rendements deviennent alors endogènes, ce ne sont plus des paramètres mais des variables des PL AROPAj.

Résultat (« sortie d'un modèle »)

Dans les modèles mobilisés, on distingue en « entrée » les paramètres et en « sortie » les résultats. Un résultat AROPAj (qu'on appelle aussi variable de résultat) est une combinaison de paramètres et de variables PL, ces dernières constituant la « solution » du PL d'optimisation.

RICA

Réseau d'Information Comptable Agricole : base de données micro-économiques fournie par un Etat membre de l'Union Européenne, et mise à disposition sous une forme standardisée par les services de la DG AGRI, sous condition stricte de l'usage. Les données sont réputées statistiquement représentatives de la production agricole à l'échelle « régionale » (les régions RICA correspondent en France aux régions administratives).

SIG

Système d'information Géographique ; une collection d'outils, normes et programmes informatiques spécifiques créée pour gérer, représenter, manipuler et analyser de manière simple et efficace toutes formes de données géo-référencées.

STICS

Modèle de simulation de culture, STICS est un modèle de fonctionnement des cultures à pas de temps journalier, développé par l'INRA depuis 1996. Son principal objectif est de simuler les conséquences des variations du milieu et du système de culture sur la production d'une parcelle agricole et sur l'environnement. A partir de la caractérisation du climat, du sol, de l'espèce cultivée et des techniques culturales appliquées, il calcule des variables de sorties relatives à la production (quantité et qualité), à l'environnement (pertes en eau et nitrates) et à l'évolution des caractéristiques du sol sous l'effet de la culture.

Lien <http://147.100.66.194/stics/>

UCS

Unité Cartographique de Sol : les UCS sont chacune décrites par leur géométrie d'une part, et par leur composition en terme d'UTS d'autre part. La géométrie d'une UCS est un ensemble de polygones

décrits chacun par leur forme et leur position géographique. Sa composition est un ensemble d'*UTS* dont on ne connaît que les parts relatives dans l'*UCS* (pour plus de détails voir le descriptif donné par GIS Sol : <http://www.gissol.fr/index.php>).

UTS

Unité Typologique de Sol : les *UTS* sont chacune décrites par une série de variables codées (environ 40 attributs) caractérisant la nature et les propriétés des sols, en surface comme en profondeur : par exemple la texture, le régime hydrique, la charge en cailloux, le matériau originel, etc.. Un dictionnaire donne la définition de chaque occurrence de variable (descriptif donné par GIS Sol : <http://www.gissol.fr/index.php>).

USTI

Unités spatiales de transformation de l'information : unités spatiales homogènes pour l'estimation des fonctions de distribution (probabilités) des « sorties AROPAj » sur un espace géographique. Ce sont des polygones (a priori non contigus) formés de cellules élémentaires et formant une partition des régions. Les transferts de l'information distribuée par AROPAj sont calculés à ce niveau, qui est le niveau le plus fin compatible avec les données économiques et physiques permettant de localiser les agents et activités de production agricole.

Variables (modèle)

De façon schématique, un *agent économique* met en œuvre une *stratégie* en déterminant des *actions* (selon les concepts prévalant en théorie des jeux). Un modèle statique comme AROPAj détermine la réaction d'un agent à un environnement donné (caractérisé par les *paramètres*, dont certains, comme ceux qui caractérisent les politiques économiques, relèvent des stratégies mises en œuvre à un autre niveau). Les actions sont les véritables variables de commande exercées par les agents économiques. Sur le plan mathématique, en programmation linéaire, l'ensemble des variables intègre évidemment ces variables de commande, et il est complété par d'autres variables qui résultent de la traduction du programme d'un agent en programme mathématique linéaire faisant apparaître des variables intermédiaires. En programmation linéaire mixte, interviennent également des variables entières. Enfin, aux variables primales que l'on vient d'évoquer (en nombre égal à celui des « colonnes » d'un programme linéaire), on ajoute les variables duales associées à chacune des contraintes du modèle (en nombre égal à celui des contraintes). Ces variables sont particulièrement importantes en économie du fait de leur interprétation. Ainsi, à une contrainte exprimant une limite sur l'utilisation d'une ressource, à l'optimum, la valeur de la variable duale associée s'interprète comme la variation de l'objectif imputable à la variation d'une unité de la ressource (on exprime cela par les expressions *prix dual*, *coût marginal*, *prix implicite* ; en anglais *dual value*, *shadow cost*, *shadow price*).

Indexation et notations

- i** indice affecté à l'activité animale (AROPAj) ou l'unité spatiale élémentaire (l'unité CLC : 100mx100m)
- j** indice affecté aux activités végétales (dont la signification varie néanmoins selon les bases de données ou les modèles : CLC, LUCAS, AROPAj, RICA)
- k** indice affecté aux groupes type AROPAj
- m** indice affecté aux mailles MODCOU
- s** indice affecté aux USTI
- p_{ij}** probabilité de présence de l'activité *j* sur la cellule *i*
- q_{ik}, q_{sk}** probabilité de présence du groupe type *k* sur la cellule *i* ou l'USTI *s*

Annexe : Groupes type V2-AROPAj des 8 régions concernant le Bassin de la Seine

On donne quelques indications sur la caractérisation des groupes type AROPAj interférant avec le Bassin de la Seine. On distingue le code RICA des Régions (voir aussi la Figure 3 pour la dénomination géographique habituelle), l'identification engendrée automatiquement par la classification automatique, et l'indice de groupe « k » utilisé pour chaque Etat Membre (ici la France).

Région	Identifiant	k									
121	101012102	1	132	101013202	13	134	101013402	26	136	101013602	40
121	201012102	2	132	102013202	14	134	102013402	27	136	102013602	41
121	301012102	3	132	103013202	15	134	103013402	28	136	103013602	42
			132	201013202	16	134	201013402	29	136	104013602	43
			132	202013202	17	134	203013402	30	136	105013602	44
131	101013102	4	132	203013202	18	134	301013402	31	136	202013602	45
131	102013102	5	132	302013202	19	134	401013402	32	136	205013602	46
131	103013102	6	132	303013202	20				136	302013602	47
131	104013102	7				135	101013502	33	136	305013602	48
131	105013102	8				135	102013502	34	136	405013602	49
131	106013102	9	133	101013302	21	135	103013502	35			
131	201013102	10	133	102013302	22	135	104013502	36	151	101015102	59
131	205013102	11	133	103013302	23	135	201013502	37	151	102015102	60
131	301013102	12	133	104013302	24	135	301013502	38	151	103015102	61
			133	105013302	25	135	401013502	39	151	202015102	62
									151	302015102	63

Tableau 6. Codes utilisés pour les différentes entités : Régions RICA et groupes type AROPAj.

Dans les tableaux suivants, on donne les indications qui fondent la classification automatique de l'échantillon RICA en groupes type AROPAj. Les valeurs réelles sont les tailles des échantillons de la base RICA.

Région 121

----- iden=201012102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
121	6	8	3.00
	13	8	56.00
	14	8	10.00
	81	8	2.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=101012102 -----

			alt	
			1	2
Region	Farm_Type	Size_Cat		
121	13	7	36.00	1.00
	14	7	1.00	.

----- iden=301012102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
121	13	5	3.00
		6	25.00
	14	2	1.00
		3	1.00
		5	1.00
		6	1.00
	81	6	1.00
	82	5	2.00

Région 131

----- iden=104013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	411	4	1.00
		5	1.00
		6	4.00
		7	1.00
		8	1.00
	412	5	3.00
		6	10.00
		7	5.00
		8	3.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=102013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	43	5	1.00
		6	5.00
		7	8.00
		8	8.00
	82	5	1.00
		8	2.00

----- iden=106013102 -----

			alt
			2
Region	Farm_Type	Size_Cat	
131	13	5	1.00
		7	7.00
		8	10.00
	43	6	2.00
		7	1.00
		8	2.00
44	8	1.00	
81		5	1.00
		6	2.00
		7	4.00
		8	5.00
411		4	1.00
		5	1.00
412		6	1.00
		7	3.00

----- iden=101013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	13	5	12.00
		6	25.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=105013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	6	7	1.00
		14	2.00
		6	6.00
		7	16.00

----- iden=103013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	81	4	1.00
		5	1.00
		6	5.00
		7	4.00
		8	10.00

----- iden=201013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	13	8	48.00

----- iden=205013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	6	8	7.00
	14	8	60.00

----- iden=301013102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
131	13	7	20.00

Région 132

----- iden=201013202 -----

			alt
			1
Region	Farm_Type	Size_Cat	
132	14	4	1.00
		5	5.00
		6	8.00
		7	20.00

----- iden=102013202 -----

			alt	
			1	
Region	Farm_Type	Size_Cat		
132	81	5	3.00	
		6	12.00	
	411	6	8.00	
		5	1.00	
	412	5	1.00	
		6	3.00	

----- iden=202013202 -----

			alt
			1
Region	Farm_Type	Size_Cat	
132	5	8	1.00
		43	2.00
		71	2.00
		81	42.00
	411	8	1.00
		412	8

----- iden=103013202 -----

			alt	
			1	2
Region	Farm_Type	Size_Cat		
132	13	3	1.00	.
		4	2.00	.
		5	7.00	.
		6	17.00	1.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=203013202 -----

			alt	
			1	2
Region	Farm_Type	Size_Cat		
132	13	8	57.00	1.00

----- iden=101013202 -----

			alt
			1
Region	Farm_Type	Size_Cat	
132	14	8	66.00

----- iden=302013202 -----

			alt
			1
Region	Farm_Type	Size_Cat	
132	81	7	12.00
	411	7	1.00
	412	7	3.00

----- iden=303013202 -----

			alt
			1
Region	Farm_Type	Size_Cat	
132	13	7	19.00

Région 133

----- iden=102013302 -----

			alt
			1
Region	Farm_Type	Size_Cat	
133	81	2	1.00
		5	2.00
		6	4.00
		7	13.00
		7	13.00
		8	32.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=105013302 -----

			alt
			1
Region	Farm_Type	Size_Cat	
133	411	4	3.00
		5	4.00
		6	2.00
		8	2.00
	412	5	4.00
		6	4.00
		7	6.00
		8	8.00

----- iden=101013302 -----

			alt	
			1	
Region	Farm_Type	Size_Cat		
133	5	7	1.00	
		42	3	1.00
			5	2.00
			7	1.00
	43	5	2.00	
		6	7.00	
		7	5.00	
		8	7.00	
	44	5	1.00	
	71	8	1.00	

----- iden=103013302 -----

			alt
			1
Region	Farm_Type	Size_Cat	
133	13	4	1.00
		5	3.00
		6	8.00
		7	10.00
		8	19.00

**

----- iden=104013302 -----

			alt
			1
Region	Farm_Type	Size_Cat	
133	14	6	4.00
		7	3.00
		8	25.00

Région 134

----- iden=101013402 -----

			alt
			1
Region	Farm_Type	Size_Cat	
134	13	8	130.00
	14	8	14.00

----- iden=203013402 -----

			alt		
			1	2	3
Region	Farm_Type	Size_Cat			
134	6	4	1.00	.	.
		5	1.00	.	.
		6	1.00	.	.
	13	3	.	1.00	.
		4	.	.	1.00
		6	.	1.00	1.00
	42	4	.	1.00	.
	81	5	2.00	1.00	1.00
		6	9.00	.	.
	411	5	2.00	.	.
	412	5	1.00	.	.

----- iden=201013402 -----

			alt
			1
Region	Farm_Type	Size_Cat	
134	13	7	64.00
	14	7	5.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

iden=102013402

			alt	
			1	2
Region	Farm_Type	Size_Cat		
134	5	5	1.00	.
		8	2.00	.
42		4	1.00	.
		5	6.00	.
		6	3.00	.
		7	1.00	.
		8	3.00	.
		43	6	1.00
44		2	1.00	.
		5	2.00	.
		6	3.00	.
		7	2.00	.
		8	1.00	1.00
71		5	.	1.00
		8	2.00	.
72		5	1.00	.
		7	1.00	.
		8	4.00	.
82		6	1.00	.
		7	4.00	.
		8	5.00	.

iden=103013402

			alt	
			1	2
Region	Farm_Type	Size_Cat		
134	6	7	3.00	1.00
		8	11.00	.
13	7	.	1.00	.
81		7	11.00	.
		8	21.00	.
411	7	4.00	.	.
412		7	2.00	.
		8	1.00	1.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=301013402 -----

			alt
			1
Region	Farm_Type	Size_Cat	
134	13	6	49.00
	14	6	4.00

----- iden=401013402 -----

			alt
			1
Region	Farm_Type	Size_Cat	
134	13	3	4.00
		4	4.00
		5	20.00
	14	2	1.00
		3	1.00
		4	1.00
		5	2.00

Région 135

----- iden=101013502 -----

			alt
			1
Region	Farm_Type	Size_Cat	
135	43	6	1.00
	411	6	29.00
	412	6	10.00

----- iden=104013502 -----

			alt
			1
Region	Farm_Type	Size_Cat	
135	6	8	2.00
	81	4	1.00
		5	4.00
		6	3.00
		7	3.00
		8	14.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=102013502 -----

			alt	
			1	
Region	Farm_Type	Size_Cat		
135	13	3	1.00	
		5	5.00	
		6	4.00	
		7	8.00	
		8	5.00	
		14	6	2.00
			7	3.00
			8	4.00

----- iden=103013502 -----

			alt	
			1	2
Region	Farm_Type	Size_Cat		
135	5	7	2.00	.
		8	1.00	.
	42	2	1.00	.
		3	3.00	.
		4	4.00	.
		5	2.00	.
71	7	1.00	.	
	8	1.00	1.00	
72	8	3.00	.	
82	7	1.00	.	
	8	1.00	.	

----- iden=301013502 -----

			alt
			1
Region	Farm_Type	Size_Cat	
135	43	8	6.00
	411	8	21.00
	412	8	9.00

----- iden=201013502 -----

			alt
			1
Region	Farm_Type	Size_Cat	
135	43	2	1.00
		5	1.00
	411	2	1.00
		3	2.00
		4	4.00
		5	20.00
	412	4	2.00
		5	6.00

----- iden=401013502 -----

			alt
			1
Region	Farm_Type	Size_Cat	
135	43	7	2.00
		411	20.00
		412	18.00

Région 136

----- iden=101013602 -----

			alt
			2
Region	Farm_Type	Size_Cat	
136	13	5	1.00
		6	1.00
		7	4.00
		8	14.00
	14	8	1.00
	81	5	1.00
		6	4.00
		7	1.00
		8	5.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=103013602 -----

			alt	
			2	3
Region	Farm_Type	Size_Cat		
136	5	8	1.00	.
	42	3	3.00	1.00
		4	4.00	1.00
		5	15.00	2.00
		6	6.00	.
		7	4.00	.
	44	5	2.00	.
6		.	1.00	
82	7	1.00	.	
	8	1.00	.	
411	6	2.00	.	
412	6	1.00	1.00	
	8	1.00	.	

----- iden=104013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	6	6	1.00
		8	4.00
	81	4	1.00
		5	3.00
		6	7.00
		7	4.00
		8	21.00
411	6	1.00	
	7	2.00	

----- iden=205013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	42	6	19.00
	44	6	2.00
	71	6	1.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=405013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	42	7	13.00
	43	7	2.00
	82	7	1.00

----- iden=105013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	42	3	1.00
		4	1.00
		5	20.00
	44	3	1.00
		4	1.00
	412	5	1.00

----- iden=302013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	13	4	1.00
		5	8.00
		6	16.00
	14	6	3.00

----- iden=305013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	42	8	12.00
	43	8	1.00
	44	8	1.00
	71	8	1.00
	72	8	1.00
	82	8	1.00

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=102013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	13	8	44.00
	14	8	5.00

----- iden=202013602 -----

			alt
			1
Region	Farm_Type	Size_Cat	
136	13	7	17.00
	14	7	2.00

Région 151

----- iden=101015102 -----

			alt	
			2	3
Region	Farm_Type	Size_Cat		
151	6	7	1.00	.
	13	6	1.00	.
		8	3.00	.
43	6	1.00	.	.
	7	1.00	.	.
	8	5.00	.	.
71	8	1.00	.	.
81	6	1.00	.	.
	7	1.00	.	.
	8	11.00	.	.
411	4	.	1.00	.
	5	1.00	1.00	.
	6	4.00	.	.
	7	3.00	1.00	.
	8	2.00	.	.
412	5	1.00	.	.
	6	1.00	.	.
	7	3.00	.	.
	8	9.00	.	.

Programme PIREN-Seine : Modélisation économique et agriculture sur le bassin de la Seine

----- iden=102015102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
151	43	8	7.00
	44	8	1.00
	71	8	2.00
	81	8	43.00
	82	8	2.00
	411	8	5.00
	412	8	12.00

----- iden=103015102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
151	6	8	2.00
	13	5	4.00
		6	8.00
		7	5.00
		8	42.00
	14	8	2.00

----- iden=302015102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
151	42	4	1.00
		6	5.00
	43	5	1.00
		6	2.00
	44	6	1.00
	81	5	4.00
		6	9.00
	411	5	2.00
		6	3.00
	412	5	1.00
		6	4.00

----- iden=202015102 -----

			alt
			1
Region	Farm_Type	Size_Cat	
151	42	7	1.00
	43	7	5.00
	44	7	1.00
	81	7	14.00
	411	7	2.00
	412	7	11.00