Mise au point d'un outil de modulation intra-parcellaire de la fertilisation azotée du blé d'hiver basé sur la télédétection et un modèle de culture

V. Houlès

▶ **To cite this version:**

V. Houlès. Mise au point d'un outil de modulation intra-parcellaire de la fertilisation azotée du blé d'hiver basé sur la télédétection et un modèle de culture. Sciences du Vivant [q-bio]. Institut National Agronomique Paris Grignon; ARVALIS - Institut du végétal [Paris], 2004. Français. NNT : . tel-02833417

HAL Id: tel-02833417
https://hal.inrae.fr/tel-02833417

Submitted on 7 Jun 2020

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

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

Institut National de la Recherche Agronomique

Unité d'Agronomie de Laon-Reims-Mons

Unité Climat-Sol-Environnement d'Avignon

Arvalis – Institut du Végétal

Institut National Agronomique Paris-Grignon

Ecole doctorale ABIES

AH TH57
Bibliothèque
I.N.R.A. - C.S.E.
Unité Climat Sol et Environnement
Bâtiment Climat
Site Agroparc
F. 84914 AVIGNON CEDEX 9
FRANCE

THÈSE

Présentée par

Vianney HOULÈS

pour obtenir le titre de

Docteur de l'Institut National Agronomique Paris-Grignon

Mise au point d'un outil de modulation intra-parcellaire de la fertilisation azotée du blé d'hiver basé sur la télédétection et un modèle de culture

soutenue publiquement le 17 mars 2004 devant le jury composé de :

P. DEBAEKE	INRA	Rapporteur
J.C. CALVET	CNRM	Rapporteur
B. NEY	INA P-G	Examinateur
A. MESSEAN	CETIOM	Examinateur
P. GATE	Arvalis – Institut du Végétal	Examinateur
M. Guérif	INRA	Examinatrice
B. Mary	INRA	Président

Directeur de thèse :
Bruno MARY, DR, INRA
Co-encadrante :
Martine GUÉRIF, DR, INRA

Index

Approche par assimilation : désigne la façon de renseigner les variables d'entrée du modèle qui consiste à utiliser les données de télédétection pour faire de l'assimilation (p. 157)

Approche par cartographie : désigne la façon de renseigner les variables d'entrée du modèle qui consiste à caractériser par des mesures et des observations pédologiques la variabilité des propriétés du sol (p. 130)

Approche indicateur : méthode de préconisation de la fertilisation azotée impliquant le calcul d'indicateur du statut azoté de la plante (p. 31)

Approche modèle : méthode de préconisation de la fertilisation azotée impliquant l'utilisation d'un modèle de culture (p. 34)

AZOBIL : méthode de préconisation de la fertilisation azotée prise comme référence (Machet *et al.*, 1990) (p. 32)

B : bilan d'azote (kg ha^{-1}) (p. 57)

Cab : teneur en chlorophylle a et b des feuilles (g cm^{-2}) (p. 37)

Critère agro-environnemental : grandeur que l'on cherche à optimiser pour choisir la dose d'azote et calculée à partir de *G*, *P*, *R* ou *B* (p. 41)

Critère A : critère agro-environnemental basé sur une taxe portant sur les bilans d'azote (p. 97)

Critère B : critère agro-environnemental basé sur une taxe portant sur les reliquats d'azote (p. 97)

Critère C : critère agro-environnemental basé sur un seuil appliqué aux bilans d'azote (p. 97)

Critère D : critère agro-environnemental basé sur un seuil appliqué aux reliquats d'azote (p. 97)

Dose bilan : dose d'azote nécessaire à la croissance de la culture calculée par la méthode de la balance azotée (*cf.* AZOBIL) (p. 32)

E : variable traduisant l'impact environnemental d'un certain niveau de fertilisation et pouvant être représentée par *R* ou *B* (p. 41)

Fonction de pédotransfert : permet de passer d'une caractérisation pédologique qualitative des sols aux entrées du modèle (p. 135)

G : rendement en grain à 0% d'humidité (t ha^{-1}) (p. 40)

Grille : dispositif de sondage par points réguliers servant à des mesures sur les deux parcelles (*cf.* recto)

GLUE : une des méthodes d'assimilation comparées ; utilisée sur les données réelles (p. 164)

GPS : Global Positioning System (p. 27)

IAP : voir information *a priori* (p. 42)

Information *a priori* (IAP) : connaissance frustre qu'on a des valeurs des variables d'entrée du modèle (p. 42)

INN : Indice de nutrition azotée (p. 32)

Krigeage : estimateur permettant de connaître la valeur d'une variable à un endroit où elle n'a pas été mesurée en fonction de mesures en d'autres points (géostatistique) (p. 36)

LAI : indice de surface foliaire (Leaf Area Index, sans dimension) ; une des variables observées servant à l'assimilation (p. 32)

M : marge brute (ha^{-1}) (p. 56)

MCIP : une des méthodes d'assimilation comparées (p. 167)

MS : matière sèche totale de la culture (t ha^{-1}) (p. 69)

N^* : dose d'azote optimale d'après les données observées (p. 41)

$\hat{N}^*$: dose d'azote optimale estimée par le modèle (p. 41)

P : teneur en protéines des grains (% de matière sèche) (p. 40)

Paramètre : (définition adoptée pour cette étude) entrée du modèle dont la valeur est constante dans le temps et dans l'espace (p. 41)

QCab : quantité de chlorophylle présente dans l'ensemble des feuilles de la culture, non simulée par STICS (kg ha^{-1}) (p. 45)

QN : quantité d'azote absorbée par la culture (kg ha^{-1}) ; une des variables observées servant à l'assimilation, simulée par STICS (p. 40)

R : reliquats d'azote dans le sol à la récolte (kg ha^{-1}) (p. 40)

RMSE : Root Mean Square Error, critère d'évaluation des simulations (p. 59)

SIG : Système d'Information Géographique (p. 28)

SIMPLEXE : nom donné à une des méthodes d'assimilation testées (p. 160)

Théorème de Bayes : exprime la dépendance entre distribution *a priori* des valeurs des variables d'entrée du modèle et leur distribution *a posteriori* grâce aux observations (p. 44)

Triangle des textures : donne la classe texturale d'un sol en fonction de la granulométrie (teneur en sables, limons et argiles) (p. 131)

UCS : Unité Cartographique de Sol (p. 132)

UTS : Unité Typologique de Sol (p. 132)

Variable d'entrée du modèle : (définition adoptée pour cette étude) entrée du modèle qui est géographiquement variable (p. 41)

Variables d'intérêt final : variable de sortie du modèle : rendement, teneur en protéines et variable environnementale (reliquats, *R* ou bilan d'azote, *B*) (p. 40)

Expérimentations du site de Chambry

Parcelle 1, 1999-2000

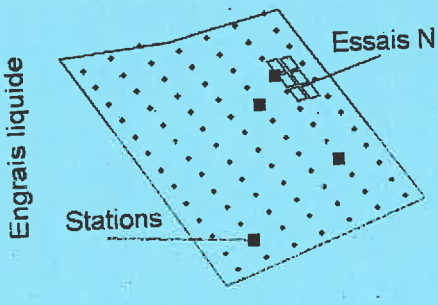
Acquisition de télédétection (CASI) : 8 avril, 6 mai, 2 juin, 28 juin

Essais doses d'azote :

0 (kg ha^{-1})	180
60	240 (parcelle)
120	300

→ Calibration des mesures de télédétection
→ Détermination des relations *QCab* → *QN*
Grille : Eau et Azote sol + observations pédologiques
Carte de rendement et des teneurs en protéines
Stations : Paramétrage de STICS, fosses

Parcelle 2, 2000-2001



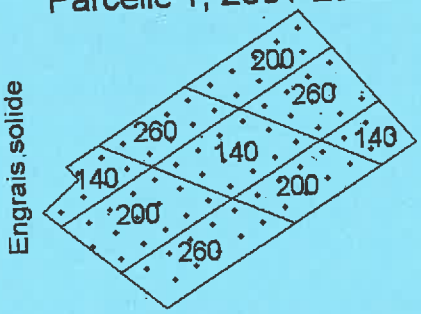
Acquisition de télédétection (CASI) : 2 avril, 21 avril, 8 mai, 31 mai

Essais doses d'azote :

0 (kg ha^{-1})	220 (parcelle)
120	260
160	280

→ Calibration des mesures de télédétection
→ Détermination des relations *QCab* → *QN*
Grille : Eau et Azote sol + observations pédologiques
Carte de rendement et des teneurs en protéines
Stations : Paramétrage de STICS, fosses

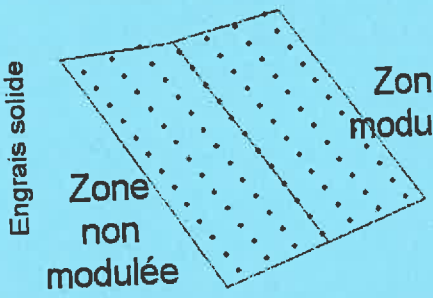
Parcelle 1, 2001-2002



Fertilisation en carré latin :
3 doses, 9 carrés

Grille : Eau et Azote sol
Carte de rendement et des teneurs en protéines

Parcelle 2, 2002-2003



Fertilisation modulée

Grille : Eau et Azote sol
Carte de rendement et des teneurs en protéines

Institut National de la Recherche Agronomique

Unité d'Agronomie de Laon-Reims-Mons

Unité Climat-Sol-Environnement d'Avignon

Arvalis – Institut du Végétal

Institut National Agronomique Paris-Grignon

Ecole doctorale ABIES

THÈSE

Présentée par

Vianney HOULÈS

pour obtenir le titre de

Docteur de l'Institut National Agronomique Paris-Grignon

Mise au point d'un outil de modulation intra-parcellaire de la fertilisation azotée du blé d'hiver basé sur la télédétection et un modèle de culture

soutenue publiquement le 17 mars 2004 devant le jury composé de :

P. DEBAEKE	INRA	Rapporteur
J.C. CALVET	CNRM	Rapporteur
B. NEY	INA P-G	Examinateur
A. MESSEAN	CETIOM	Examinateur
P. GATE	Arvalis – Institut du Végétal	Examinateur
M. Guérif	INRA	Examinatrice
B. Mary	INRA	Président

Directeur de thèse :
Bruno MARY, DR, INRA
Co-encadrante :
Martine GUÉRIF, DR, INRA

A mos pepin e menina pairals, aquels desconeguts, boratièrs de la montanha tarnesa, desirant qu'aguèssen pas trapat tot çò que seguirà per tròp sobrenc.

A mes grands-parents paternels, ces inconnus, métayers de la montagne tarnaise ; en espérant qu'ils n'eussent pas trouvé tout ce qui va suivre par trop superfétatoire.

Aimablement traduit en « Languedocien » par L. Thuriès

Chaque fois que la science avance d'un pas,
c'est qu'un imbécile la pousse, sans le faire exprès.
Emile Zola

Remerciements, avant-propos, exergue oni(ri)que

Pense au bûcher funèbre et, quand tu le peux encore, mêle un peu de folie à tes desseins : il est doux de déraisonner à propos.

Horace, Odes, IV, 12-17 (Ça fait toujours bien de citer les vers d'Horace[1])

Rien n'est plus sot que de traiter avec sérieux de choses frivoles ; mais rien n'est plus spirituel que de faire servir les frivolités à des choses sérieuses.

Erasme, Eloge de la Folie (Je ne suis pas sûr d'avoir compris, mais ça ressemble à une devise Shadok, ce qui prouve bien qu'Erasme était en avance sur son temps)

Puisse cet espace de libre expression, c'est du moins ainsi que je le perçois, trouvant sa place avant la description circonstanciée de travaux qui ne donnent guère d'occasion à des digressions oscillant entre des considérations relatives au caractère bucolique des égaillements[2] futiles mais nécessaires dans les folles prairies de l'insouciance comme dirait l'autre, pas plus qu'à des rêveries mélancoliques relatives à la brièveté de notre passage sur terre et à l'imminence de la dégradation, autrement dit et pour user d'un terme agronomique (faut bien s'y faire), « minéralisation », de notre enveloppe organique qui ne saurait durer plus longtemps à l'encontre de certaines piles, me permettre de remercier comme il se doit l'ensemble des personnes qui m'ont aidé, sur le plan scientifique ou amical, à mener (à bien ???) iceux travaux champêtres certes mais guère romantiques donc. Tant, dès lors, qu'un sybaritisme verbal me sera permis et avant qu'il ne doive, bien malgré lui, céder la place à un style d'expression anachorétique pour ne pas dire ascétique afin de ne confondre point avec l'acide du presque même nom, tant qu'on me permettra ce style pompeux et pompant j'en conviens mais ô combien salvateur, veuillent les Piérides dont les noms m'échappent pour la plupart et après les insignes sacrés desquelles je cours en vain dans la folie qui m'égare, accueillir mon fromage et m'inspirer les différents synonymes de remerciements que cèle le français idiome (voyons, voyons : gratitude, reconnaissance, savoir gré à, être l'obligé de, être redevable de, déférence, égard, plus limite ; hum, ça va pas faire lourd cette histoire ; il y en eût eu sans doute davantage en Sanskrit, oui mais voilà, je ne le pratique pour ainsi dire pas du tout). Oh ! Où sont les synonymes ?[3]

Dans un premier temps donc, que Calliope et Polymnie (Robert (le petit) m'a fort sympathiquement rappelé les noms des Muses ainsi que leurs fonctions respectives, merci à Robert (à le petit)) guident la bille de mon stylo bic (ça en jette moins que les envolées lyriques avec la plume des poètes mais que voulez-vous, il faut bien vivre avec son temps et oui ma p'tite dame et avec ça qu'est-ce que je vous sers, ce sera tout merci, il y en a un peu plus, je vous le mets ? oui, bien sûr, on n'en est pas à 200 g de matière organique décédée près, vous avez vu il s'est mis à pleuvoir ah ça y a plus de saison ah là là, où qu'on va je vous le demande, et encore j'eusse pu directement dactylographier la présente, mais que voulez-vous, c'est mon petit côté rétrograde), puis que Thalie

---

[1] Mais pas ses asticots, sinon, vous voyez ça d'ici : Asticot n°1 : « J'ai mangé le nez d'Horace, hmm, bof, un peu dur ; s'il eût été plus court, cela n'eût pas changé la face du menu » ; Asticot n°2 : « Moi la cuisse haute, il avait le jarret vigoureux, le père Horace, j'ai eu du mal à en venir à bout mais j'y suis parvenu, fatigué mais heureux, et j'essaierai de faire mieux la prochaine fois » ; Asticot n°3 : « J'ai écopé du pied, il avait de très bons pieds ailés, Horace aux grands pieds » ; Asticot n°4 : « Et moi du cerveau, très fin, très tendre, très relevé : quel grand poète est mort avec toi, Horace, ô désespoir » ; ça m'asticote cette histoire.

[2] De : égailler.

[3] Le lecteur attentif aura reconnu en cet affreux salmigondis un outrancier et ostensible plagiat, aux parenthèses près, des Géorgiques, II, 475 et suivants peut-être.

et Nimportekhoi (cette dernière n'étant pas avérée par Robert (le petit), sacré Robert (sacré le petit)) prennent le relais.

***

En premier lieu, que s'exprime toute ma gratitude envers mes directeurs/encadrants de thèse (je n'ai jamais très bien compris la différence, ce doit être d'ordre administratif), Martine Guérif et Bruno Mary, notamment pour le recul sur le sujet qu'ils m'ont permis de trouver au travers de longues et toujours enrichissantes conversations que nous avons pu avoir sur celui-ci, pour leur engagement dans mon travail, leur attention et leur rigueur. Je dois souligner que j'ai eu la chance d'être encadré par deux directeurs de recherche qui de surcroît se connaissent depuis de nombreuses années et s'entendent bien, ce qui est un facteur important pour la réussite sereine d'un projet de ce genre. Si jamais la recherche devait devenir mon gagne-pain (et par les temps qui courent, il est permis d'en douter, non que je ne veuille pas de la recherche, mais qu'elle ne veuille pas de moi), il est probable que leur influence, d'une façon ou d'une autre, se ressentira sur mes activités pendant quelques années. Je ne peux qu'avoir un mot particulier pour Martine, qui m'a particulièrement aidé dans la seconde moitié de ma thèse, notamment au cours de la rédaction.

Je saurai également gré aux directeurs des unités qui m'ont accueilli (il y en a eu deux, je n'ai pas voulu abuser et imposer mon importune présence par trop longtemps) : Sylvie Recous à l'unité d'Agronomie de Laon-Reims-Mons à Laon, pour la première moitié de la thèse, dont je ne peux que saluer les qualités humaines qui l'ont conduite, entre autres, à se soucier des modalités pratiques de mon déménagement (alors qu'avec la fusion de Laon-Reims-Mons, n'est-ce pas, il y avait fort à faire) ; André Chanzy à l'unité Climat, Sol et Environnement en Avignon, pour la seconde moitié de la thèse, qui a bien voulu que je reste plus longtemps que prévu et qui n'a pas rechigné à me doter d'un puissant ordinateur au bon endroit et au bon moment comme il est d'usage de dire en agriculture de précision et ce en pleine crise des 30%.

Cette thèse a bénéficié d'un cofinancement INRA-Département Environnement & Agronomie et Arvalis, Institut du Végétal (anciennement ITCF, Institut Technique des Céréales et Fourrages : que de changements au cours de cette thèse : le passage à l'Euro, à Arvalis...). Je saisis l'occasion pour remercier les représentants de cet organisme et plus particulièrement Philippe Gate et François Laurent, qui ont collaboré à ce travail de diverses manières, tant par la fourniture ou l'acquisition de données que par leur présence constructive aux comités de pilotage. J'associe volontiers à ces remerciements Gérard Aubrion et Sébastien Vautrin, de l'ITCF également, pour leur coopération sur les données de 2002, ainsi que toutes les autres personnes qui ont participé aux comités de pilotage, dont les noms reviendront par la suite mais dont la liste exhaustive suit céans : Frédéric Baret, Jean-Paul Bordes, Nadine Brisson, Philippe Gate, Eric Justes, François Laurent, Jean-Marie Machet, David Makowski, Sophie Moulin, Bertrand Ney, Hervé Poilvé.

Je serai également redevable aux personnes qui ont fait partie de mon jury de thèse pour leurs remarques et questions intéressantes, Jean-Christophe Calvet et Philippe Debaeke comme rapporteurs, et Philippe Gate, Antoine Messean comme examinateurs, outre mes encadrants. Je remercie également Françoise Launay, de l'école doctorale ABIES, pour sa serviabilité et sa patience pour le traitement des tâches administratives.

Le projet « Agriculture de Précision » au sein duquel s'inscrit le sujet du présent travail et initié par Martine Guérif, s'appuyant sur une expérimentation située non loin de Laon, fut l'occasion pour quatre unités INRA de travailler ensemble : l'unité d'Agronomie de Laon-Reims-Mons, l'unité de science du sol d'Orléans et les unités CSE et de biométrie d'Avignon, collaboration dont j'ai largement été bénéficiaire. Que s'exprime donc mon égard envers Bernard Nicoullaud qui n'a pas ménagé sa peine pour faire les cartes pédologiques de Chambry, Alain Couturier qui s'occupait des SIG, Claude Bruchou pour la spatialisation des données, Denis Allard qui m'a suggéré une façon de montrer que ce que simule STICS n'est pas tout à fait aléatoire, Sophie Moulin (comment ça elle est trop blanche la moquette maintenant ?), chargée de recherche et de l'inversion de données de télédétection, pour ses nombreux conseils, particulièrement à mes débuts tumultueux avec Matlab, Nicolas Beaudoin, autre pédologue dans l'âme, thésard au grand cœur et fidèle à ses convictions (c'est promis, j'ai encore espoir de trouver une solution technique pour débarrasser un jour votre grenier de mon importun fourniment) et Jean-Marie Machet, entre autres pour les longues discussions que nous avons pu avoir le soir, alors que peu à peu l'antique antre de science (anciennement hôpital psychiatrique crois-je savoir) axonnais se vidait de ses pensionnaires, et qu'une douce pluie d'automne (même en plein été, c'est ça qui est fort) se déversait du ciel obscurci, abandonné de l'astre du jour se parant de son ensemble cotonneux gris diapré de purpurin, sur le rougeoiement chamarré de la vigne sauvage s'étirant sur les murs du susdit antre de science

dendritique[4]. Enfin, je ne saurai que trop remercier les deux post-doctorants de la dernière heure (l'effet deux en un) qui ont bien mérité le talent d'or de la société générale pour leur engagement, leur rigueur et leur discipline (pour cette fine allusion rugbystique (désolé Matthew (il aime pas le rugby)), *cf. infra* (va falloir vous y habituer à ces *cf. infra*)) : Matthew Pringle, qui a mis bon ordre dans la base de données en moins de temps qu'il n'en faut pour le dire (il faut que je mette quelques mots en Anglais, là, c'est quand même la moindre des choses), thank you for your efficiency which simplified my work at the end, all the corrections of my English you have made and your australian touch during the 2003 summer ; ça devrait suffire, et Clement Atzberger qui l'a remplacé au pied levé, pour son apport dans la réflexion finale et les perspectives, ces fameuses petites choses où l'on met tout ce que l'on aurait bien voulu faire sans avoir jamais trouvé le temps de le faire ; bon, comme je me suis fendu de quelques mots en Anglais, je me dois bien d'en faire autant en Allemand (parce qu'il est Allemand, l'individu susdit), l'apprentissage laborieux de cette langue m'aura au moins servi à quelque chose, mais, ouh là, ça va pas être facile, Vielen Dank für deine Arbeit und für dein Fairplay an den Brettspielen und Entschuldigung daß, ich kein Schornsteinfeger gefunden habe, Donnerwetter, ça y est, c'est fait, et, à présent, pour faire mon parfait cuistre, alors même que pas un post-doctorant grec ne vint poser le pied de sa main chez nous, surtout un Grec ancien, il ne me manquerait plus qu'à mettre une phrase en Grec, ce que je fais céans, Ο λευκος καλος ιπποσ τακεως τρεκει, le beau cheval blanc court vite pour ceux que ce genre de chose pourrait intéresser.

Tout cette belle aventure qui n'est, comme vous l'allez constater, pas sans rappeler l'épopée de Simbad le Marin, n'eût pas été possible sans la coopération obligeante de deux agriculteurs axonnais, héritiers des Templiers de la ferme de Puisieux, près de Chambry, près de Laon, dans l'Aisne, donc : Jacques et François Fontaine auquel j'associerai sa femme Jihane pour son hospitalité. Ils nous ont livré deux de leurs parcelles avec enthousiasme et intérêt ; même le coup des moquettes ne les a pas rebuté. A ce propos, je joindrai à ces remerciements Christel Noizet, plus connu sous le sobriquet de « le pilote », pour son flegme et sa coopérativité, ainsi que pour les quelques vingt minutes d'hélicoptère passées au-dessus de la campagne Picarde (très jolie) et pour les occasions qu'il nous a données de voir, sans consommer de substance illicites, un écureuil bleu volant en crabe dans l'azur picard. Je remercie également bien naturellement Rémi et Bertrand, ouvriers chez MM. Fontaine, pour leur gentillesse.

Au cours de ces trois années, j'ai eu la chance de pouvoir travailler ou avoir des contacts avec des chercheurs n'appartenant pas au projet Agriculture de Précision de Laon. Par ordre chronologique, et thématique, j'exprimerai ma reconnaissance envers Jacques Le Gouis pour ses données « chlorophylle-azote » sur d'autres variétés, Erwin Dreyer dont les conseils et les références bibliographiques fournies avec sollicitude m'ont été précieux, Eric Justes pour son aide sur STICS et la relecture de ce manuscrit, Daniel Wallach et Marie Tremblay qui m'ont accordé la faveur d'une audience à propos de l'estimation des paramètres d'entrée d'un modèle, David Makowski qui m'a beaucoup aidé tant pour le choix des critères agro-environnementaux que pour l'assimilation de données dans les modèles de culture, sans compter ses relectures critiques de plusieurs manuscrits, François Laurent et Philippe Gate pour les raisons évoquées ci-dessus, Françoise Ruget pour son aide sur STICS et sa relecture de la thèse, Jérôme Demarty évidemment sur lequel je reviendrai dans quelques lignes, pour son apport d'une méthode tellement originale qu'elle suscite des débats non consensuels et grâce auquel j'ai pu mener à terme (évitons l'expression mener à bien) cette thèse, puisque aussi bien il m'a sauvé un jour la vie, alors que nous allions à la cantine devisant de choses et d'autres et que je dérapais fort maladroitement sur une plaque de verglas (quelle fin ridicule ne m'a-t-il pas alors évité, je vois d'ici l'épitaphe) ; c'est que j'eusse fort bien pu me fracasser la crâne avec tout ça, avec toutes les conséquences que cela eût pu avoir (pour les dites conséquences, je vous renverrai aux forts belles et touchantes descriptions qu'a pu faire Homère de cervelles se répandant à terre dans ses deux best sellers, parce que là j'ai vraiment pas le temps), Philippe Clastre aux commandes de son SIG (parce que moi, sorti de Matlab, je connais plus rien) et, naturellement, Nadine Brisson, démiurge de ce monde parallèle qu'est STICS, avec ses faunes, ses nymphes, ses sylphes, ses mystères valant bien ceux d'Éleusis, pour sa disponibilité, et Dominique Ripoche, sorte de psychopompe de cet univers énigmatique (psychopompe, oserai-je le préciser, ne signifie pas qu'elle m'a pompé l'air, c'est tout le contraire, et ça n'a rien à voir avec les Shadoks non plus), pour ses modifications du code (de la matrice ?) plus nombreuses et réitérées que ce que j'escomptai, sans oublier la recette du pain d'épices.

***

(Les trois petits astérisques, ça fait bien, ça fait littéraire)

---

[4] Je sais, ça devient déjà n'importe quoi, il faut resserrer les boulons quelques lignes encore. Dendrite et axone appartiennent au champ lexical du neurone et les habitants de l'Aisne s'appellent les Axonnais voyez-vous, allez savoir pourquoi, quelle drôle d'idée vraiment ils ont eu là.

La première moitié de ma thèse s'est déroulée à Laon, fière cité picarde aux deux cœurs, dite montagne couronnée avec un à propos qui en laissera plus d'un perplexe, à la cathédrale gothique du haut de laquelle quarante bœufs nous contemplent de leur regard bovin mais pour des bœufs ce doit être normal, cathédrale à la nef élancée et risquant de s'écrouler, dont les poètes n'ont de cesse de chanter les louanges, bourgade aux multiples abbayes (un peu en ruines quand même), à la chapelle des Templiers, terre promise des Poilus du Plateau de Californie, longtemps désirée et jamais atteinte pour beaucoup. C'était aussi accessoirement et plus modestement le fief de l'expérimentation agriculture de précision de Chambry, aussi remercierai-je en premier lieu l'ensemble de la joyeuse équipe des techniciens de Laon qui a grandement participé à icelle, avec, par ordre d'entrée en scène : au CIMEL, aux moquettes, au LICOR, à la station météo, au broyage des végétaux secs, au GATOR, à l'organisation générale des expéditions, et j'en passe : Frédéric Mahu ; au N-tester, à la vivisection du blé et au pâté de lapin : Daniel Boitez ; à la perforeuse et au parapluie : Caroline Dominiarczyk (ah, ces trous) ; à la caméra Xybion et au copilotage de Christel Noizet : Frédéric Bornet ; aux fosses (pédologiques, pas communes), au GATOR et au N-tester entre autres : Eric Venet ; au découpage du blé, au broyage et aux coups de mains spontanés : Patrick Devaux ; au GATOR et au N-tester (et à la barbe également) : Eric Gréhan[5] ; au découpage nostalgique des épis de blé et à l'entretien des voitures : Yannick Geoffriau ; au N-tester, au parapluie et aux boulettes : Laurence (ou Florence ?) Foulon ; au N-tester, au parapluie et aux boulettes : Florence (ou Laurence ? Mireille, c'est plus simple) Barrois ; aux boulettes : Muguette Boucher ; aux analyses de nitrates : Gaëtan Bentivegna ; aux analyses d'azote total : Olivier Delfosse (auteur d'intéressantes réflexions sur la nature de la matière), sans oublier Jérôme Duval et les Rémois : Marie-Jeanne et Christian Herre (donc : « Un sot portait un sceau dans un seau peu large. Le sot trébuche : l'étroit seau tombe. »), Sylvie et Francis Millon, Pascal Thiebeau et Gonzague Alavoine pour leur sérieux appui. Je n'oublierai pas non plus Brigitte Rocourt qui a volé à mon secours lorsque, étourdi, je me retrouvai sans papier en Espagne, et sans château non plus, surtout de cartes donc, tremblant à l'instar de Sancho Pansa de tomber sur la Sainte-Hermandad à tous les coins de rue et Valérie Dazin à laquelle je puis à présent prêter un serment sur l'honneur que plus jamais (hélas, ô temps suspends ton vol) je ne ferai sécher de saucisse aux chambres des stagiaires, pour leur disponibilité pour résoudre les tâches administratives que je ne t'y comprends rien à tous ces trucs-là, j'aurai même pu ne plus être payé en troisième année sans leur célérité expresse, voyez-vous ça, on vit dans un monde, tout de même. Je saluerai l'avenir de l'unité, les baroudeurs africains Patricia Garnier et Joël Léonard (un de mes gourous Matlab) et Pauline Defossez qui a bien failli abriter mes ustensiles pour une durée indéterminée (*cf.* Nicolas Beaudoin ; quoiqu'il paraît que *cf.* ne se met pas en italique, allez savoir) et les chercheurs que je n'ai pas déjà cités (j'espère n'en pas oublier, il y en a tellement) : Guy Richard (je n'ai pas pu recaser lors du simulacre de soutenance (je dis simulacre parce qu'on ne m'a pas revêtu d'une histrionesque simarre ni même chaussé la tête d'un bouffonesque chapeau à la fin. Pas de simagrées, donc ; il est vrai que six magrets, ça fait beaucoup. C'est pas pour dire, mais en France, depuis qu'on a coupé celle (de tête) de quelques nobles, on n'a plus guère le sens du protocole et du clinquant. Cela dit, j'ai horreur des chapeaux et du cérémonial pompeux, alors de quoi me plains-je, je vous le demande un peu) le parallèle entre l'apprentissage immémorial des propriétés des parcelles par les agriculteurs et celui par le fameux outil mais je le retiens), Yves Duval, grand amateur de photographie, Pascal Dubrulle, fier Picard à l'humour cinglant, Carolyne Dürr qui, par sa conversation, m'a permis de mieux comprendre le fonctionnement du petit monde de la recherche agronomique et ce même si elle a déclaré un jour, comme ça, tout cru, sans crier gare, que j'avais, je cite non sans un certain dégoût, « l'humour pipi caca », alors même qu'il eût été de bien meilleur aloi d'user des termes miction-fèces, beaucoup mieux perçus de l'Académie Française à ce qu'il se dit auprès des personnes qui s'y connaissent, et alors même que par ailleurs c'est absolument faux, à la limite de la tentative éhontée de diffamation calomniatrice, nous en conviendrons tous aisément (par exemple, il ne serait pas dans mes façons d'écrire dans des remerciements de thèse qu'un des plus beaux jours de ma vie (bien morne au demeurant il est vrai) fut celui où j'appris par l'entremise de Robert (le petit), l'existence de William Prout, éminent scientifique par trop oublié de l'ingrate populace et qui (et là je cite Robert (le petit toujours, n'ayons pas la folie des grandeurs)) « émit l'hypothèse selon laquelle les masses des éléments chimiques représenteraient des multiples entiers de celle de l'hydrogène », ce n'est pas rien, fallait le faire, j'y aurai jamais pensé tout seul personnellement, surtout en 1815 où il s'en fallait encore d'un bout de temps que je sois né, et même si j'eus été né alors, j'eusse peut-être été en ce moment-là même à Waterloo en train de crier Merdre aux fieffés Anglais (ce qu'ils étaient grossiers à l'époque) avant de me faire estourbir par iceux, décidément bien peu respectueux de notre grandeur, et en outre, non content de lui (là je reparle de Marcel Prout ; pardon, William, Marcel c'est son frère mais il n'a rien fait de spécial, faut dire qu'avec un prénom pareil, ça handicape), ne voilà-t-il pas qu'il ne

[5] Auteur du fameux et quotidien : « Un petit café… dans sa culotte ? », et du non moins inénarrable et mémorable (faut dire qu'à force de l'entendre tous les midis ça peut pas s'oublier) : « Il vaut mieux un petit café qu'un grand constipé », sans oublier le : « Les plombs ont sauté chez toi ? ». Je préconiserai cependant le cas échéant, non pas les bains de siège à l'instar de Rica Zaraï (source : Coluche), mais de citer vos sources, à savoir la plupart du temps Coluche, comme il sied à tout bon scientifique (je dis ça mais je ne le fais pas toujours, mais après tout je n'ai jamais prétendu être un scientifique).

s'est pas arrêté là et qu'il nous a établi la présence d'acide chlorhydrique dans le suc gastrique. Franchement, je sais pas pour vous mais moi, ça me scie ça. Il ne me viendrait jamais non plus à l'idée d'écrire dans des remarques liminaires de thèse qu'un autre des plus beaux jours de ma vie fut celui où, revenant du théâtre en compagnie de joyeux drilles que je ne citerai pas ici afin de ne les point compromettre, nous en vînmes à mettre au point le calembour que voici. Prague, 1883. D'une fenêtre tardivement éclairé, on entend le cri qui suit : « Oh, Mme Kafka, c'est un bien joli Kafka que vous nous avez fait là ». Cela se passe de commentaire j'en conviens, si ce n'est que, comme l'a paraît-il insinué Hugo (18??) dans une phrase citée par Desproges (19??), le calembour n'est jamais qu'un pet de l'esprout, pardon, de l'esprit), Jean Boiffin, qui, malgré ses hautes fonctions hiérarchiques ne manquait pas de me prodiguer affablement ses conseils et avis au détour d'une rencontre fortuite au coin d'un étroit et torve couloir laonnois, même s'il tenait mordicus à ce que je fusse le petit fils spirituel de je ne sais qui, alors que je ne suis jamais que celui de Noël et Maria Grand d'une part et de Emile et Marie Houlès d'autre part, ce qui est déjà beaucoup me semble-t-il, en tout cas ça me suffit. J'adresserai un grand salut à tous les stagiaires et assimilés (ceux que d'aucuns appellent, en ces temps troublés où le loup est de retour dans nos campagnes et où le fer des charrues est battu pour forger des glaives, les intermittents de la recherche) rencontrés à l'INRA de Laon, voire à l'anciennement SAA et nouvellement je ne sais plus ils sont pénibles de changer de nom, tous, en tout cas à l'époque c'était toujours la SAA (Station Agronomique de l'Aisne, et pas Société Anonyme des Alcooliques, faut-il le préciser), Jean-Baptiste Delerue, survivant de l'été 2000 qui avait rempilé à l'INRA pour tâter de la motte de terre avant de disparaître corps et âme dans le monde austère et fort peu champêtre de la finance, Laurent Pazek, qui a participé au mémorable aménagement des chambres neuves, fier sanglier ardennais, Samuel Abiven, vrai chercheur en herbe, aficionados de bibliographie scientifique (c'est dire), féru de littérature et d'ail, pas peu fier de m'avoir vaincu à Hannibal (mais il faut dire que je n'avais vraiment pas eu de chance cette fois-là, oui, je sais, il faut savoir la gérer, mais cette fois-là je ne pouvais rien faire, d'ailleurs j'ai vu des parties aussi dictées par l'aléa depuis, enfin bref, on ne va pas remettre ça), pour sa conversation éclairée sur maints sujets (d'ailleurs je comprends pas toujours tout) ; j'accepterai volontiers tes excuses faites au nom du peuple breton envers le mien et contre tout pour avoir participé à la Croisade contre icelui, c'est bon pour une fois mais qu'on ne vous y reprenne plus, et puis d'ailleurs je n'y étais pas du temps de la Croisade (par contre, je n'ai toujours pas saisi cette histoire d'arracheur de navet qui montre le chemin avec un navet. C'est normal, on fait avec ce qu'on a, il allait pas le montrer avec une dent, le mec. Et puis, est-ce que ça marche aussi avec les choux ?), Nawal Ezzine, native de « Casa », pour ses éclairements sur la culture marocaine, Olivier Ancelin pour ces civets de sanglier (pas celui des Ardennes, mais de la Somme) hebdomadaires (ça réconcilierait avec la chasse ces choses-là), Eva Pohl, petit farfadet germanique montée d'Avignon pour la récolte de betterave (impressionnant, la récolte de betterave, faut voir ça avant de mourir, c'est comme pour le clocher de Rodez), Viviane Simard, débarquée à Laon à la saison des barbecues pour une durée de six mois et qui y est toujours restée (ça me rappelle quelqu'un), Joachim Pereira et sa famille qui ont passé quelques mois bien loin des cocotiers brésiliens (il y a des cocotiers au Brésil ?), Christelle Perretti (« ma » première stagiaire, ça compte), Camille Ricochon, venue en renfort pédologique, Jorge Hernandez, uruguayen qui a supporté avec flegme que personne ne sache prononcer correctement son nom (faut dire que c'était pas facile), Olivier Gomez qui a hâté le dépouillement de la récolte 2002, et ceux de la fin de ma période laonnoise, les derniers avec lesquels j'ai pu profiter des bars laonnois qui servent de la vraie bière, Pierre-Vincent Protin, Grégory Dhellemme et Bertrand Jauffret, petit gars bien de Marseille fourvoyé en Picardie et égaré à Toulouse ces derniers temps. J'aurai un salut tout particulier pour les thésards de Laon, Karima Khalil, celle qui a failli ne pas venir d'Avignon et que j'ai lâchement abandonnée à la fin de sa thèse pour y aller, aux éclats de rire si particuliers, amatrice de villes aux larges rues avec des gens qui bougent dedans et qui m'a donné un éclairage différent sur la culture marocaine, ce qui ne l'a néanmoins pas empêché d'accepter de bonne grâce le fait que je portasse des pantoufles au bureau, et Filip Coppens, représentant de ce peuple dont Jules César a dit qu'il était le plus brave de tous les peuples gaulois, pour son humour tout belge (donc), comme les frites, et Dieu sait que ce n'est pas facile de plaisanter dans une langue étrangère (désolé de t'avoir laissé tomber vieux). Je n'oublierai pas non plus le cuisinier de la cantine de Laon et le reste du personnel, parce que s'il est bien une catégorie socioprofessionnelle précieuse, c'est celle des cuistots, et, pour réparer une vieille faute, Eva Kamphorst que j'avais complètement oubliée dans mes remerciements de DEA, allez savoir pourquoi, comme quoi ce sont des choses qui arrivent, alors même qu'elle avait toujours des biscuits dans son bureau, tradition qui s'est bien perdue hélas depuis, allez savoir pourquoi aussi. Enfin, pour clore cet épisode, j'aurai une pensée pour André Marin-Laflèche, auteur du désormais célèbre aphorisme qui mériterait la place du Gnothi Seauton sur le fronton d'un certain type d'écoles : « L'agronomie n'est pas une science, mais un art de vivre ». A méditer, même si ce n'est pas du Laporte (*cf. infra*).

A tous les Laonnois je dis donc : « Jamais je ne mangerai de Maroilles sans penser à vous, beuuuh uuuuiiii hêêêêê » (c'est la façon picarde de prononcer : « Si fait », malheureusement très difficilement restituable à l'écrit).

***

A ce stade, permettez-moi de m'écrier : « L'approche heuristique n'est-elle pas éristique ? ». Ça faisait un bout de temps que ça me brûlait les lèvres. La question se pose, aussi le fais-je, vu qu'elle est fondamentale pour la compréhension du monde actuel (mais si, mais si, je vous assure). Pourtant, assez peu se lèvent hors des rangs afin d'essayer d'y répondre, ne fût-ce que par bribes, mais nous savons tous que les petites bribes font les grandes découvertes. Cela montre assez à mon sens la désolidarisation des générations actuelles des vrais problèmes de fond, comme de savoir si l'approche heuristique ne serait pas éristique par le plus grand des hasards. Ceci me permet, avec un sens de la transition qui m'est particulier, d'introduire la citation libre, exécutée sous forme de florilège, de propos plus ou moins ouïs de la bouche de Bernard Laporte, et qui est indispensable (i) vu le contexte actuel et (ii) à la bonne intelligence de ce qui va suivre : « Il faut respecter les fondamentaux : de la rigueur, de la discipline, de l'engagement dans les regroupements. Pas de faute. Pas de faute. Il faut faire vivre le cuir, créer le mouvement, les ouvertures, relever le défit physique proposé par l'adversaire, ne pas reculer devant l'impack. On va les taper ». A méditer dans les chaumières, paroles de sage.

***

Sur ces bonnes paroles pleines (i) de poésie aux tendres accents enfantins et enjoués mêlés de longs sanglots monotones et (ii) de foi et d'espoir en l'avenir obscur et des lendemains qui déchantent, nous voici donc parvenus, mine de rien, comme le temps passe, c'est fou ça, à la partie avignonnaise de ma thèse : Avignon, cité des papes, du pont du même nom ou Saint-Bénézet pour les férus de détails superfétatoires, détruit en 1226 par ce fieffé Louis VIII allant guerroyer en terre languedocienne (affublé de quelques Bretons sans doute), ce dont il fut mal inspiré puisque aussi bien il se vit pris en chemin de coliques n'ayant rien de bucolique et qui l'ont obligé à renoncer à ses projets belliqueux pour cause de mort prématurée, paix à son âme mais franchement on ne pourra pas s'empêcher de penser qu'il l'avait un peu cherché. Avignon, donc, son pont, son festival, bref, vous voyez d'ici le tableau, je ne vais pas vous en esquisser un croquis, terre où souffle le Mistral et défèquent à foison les individus de race canine pour plagier plusieurs personnes en même temps, à vous de retrouver.

Dans un certain désordre spatio-temporel, je remercierai donc outre les personnes que j'ai déjà citées ci-dessus par-ci par-là, prenez des notes si vous êtes perdus, Jacques et Jacqueline Saumade, qui a fait un pot de départ qui a placé la barre des pots de départ à une hauteur telle que je ne saurai la relever, Nadine Bruguier et Véronique Desfonds, organisatrices de ballades estivales, Liu Thon, qui semble avoir réussi là où Gilgamesh a échoué, Gilbert Sappe, l'homme-station-météo, Jean-François Hanocq, le pilote-maison, Patrice Lecharpentier, de grande ressource informatique, Roland Bosseno, mon co-bureau, qui ne s'est jamais plaint lui non plus de mon port de pantoufles (alors que c'est un baroudeur, lui, pas un pantouflard) bien que je le soupçonne d'avoir été un espion à la solde des pédologues puisqu'il est pédologue (ils m'ont suivi partout), Dominique Courault (promis, je n'inverserai plus les boutons des portables, c'est comme pour la saucisse, c'est fini tout ça), Marie Weiss que son affection pour les zèbres (moi c'est plutôt les ours) n'a pas empêché de m'encourager quotidiennement de la façon la plus délicate qui se puisse imaginer à persévérer dans mon travail (« Alors, t'as pas un truc à faire, là peut-être ? »), Albert Olioso, qui n'est donc pas, vous l'aurez compris ou c'est à se taper la tête contre les murs et ma foi, non merci, vraiment, sans façon, faites le sans moi, qui n'est donc pas mon gourou en terme du port de pantoufles au travail, comme cela a pu être dit par des jaloux, étant donné qu'à Laon déjà j'en arborai fièrement et que tout petit déjà je partais souvent à l'école en oubliant de me chausser, c'est dire, tout au plus peut-on parler de convergence adaptative et d'ailleurs je ne porte pas de pantalons troués, Frédéric Baret, un coureur de fond qui entraîne dans son sillage tout un tas de gens qui du coup renâclent à faire du sprint (ce serait un gourou lui que ça m'étonnerait pas plus que ça), M. Frédéric Huard (il aime bien qu'on l'appelle monsieur), invité surprise et des derniers mois des JdPx, Xingfa Gu, qui assure un échange fertile de compétences sino-française (ou franco-chinoise, c'est selon), Odile Lusa et Roseline Olivier qui, à l'instar de leurs homologues picardes, m'ont sauvé la vie et assuré la pitance plus d'une fois, Olivier Marloie, encore un picard reconverti en provençal, et qui n'hésite pas à faire la promotion de son ancienne patrie avec une verve qu'on n'avait plus guère ouïe depuis Virgile ou Elrond, Marie Launay, encore une picarde reconvertie en provençale, il y en a plein je vous dis, et qui est en quelque sorte ma grande sœur de thèse, même si elle fait du S+ et qui a fait un plongeon dans le Styx, plouf, j'espère que tu ne te noieras pas, Dalila Mohrath et Sylvain Barbace, des « gens de sol » rencontrés dans le bus, comme quoi, l'avenir de CSE est dans le bus et il y a aussi des gens très bien dans ce bâtiment. J'en arrive à mes commensaux réguliers ou occasionnels, à la foule des instables de la recherche (comment dit-on déjà ? les précaires plutôt) ; au tout début du séjour, Diane de Pas et Emmanuel Abadie, un gars qu'était dans mon école quand j'étais jeune figurez-vous, voyez-vous ça, que le monde est petit, hein, bon, pour leur respect des fondamentaux, leur rigueur, leur discipline, leur engagement dans les regroupements, Tom Wassenaar, polyglotte qui est parti un peu vite mais qui avait l'air marrant, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Yu Tao, amateur éclairé de formes callipyges, une vieille connaissance qui avait à l'époque essayé de m'apprendre un jeu de carte chinois (mais il était pas d'accord avec un autre Chinois sur les règles et j'ai rien compris), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Mehrez Samaali,

qui, par une nuit de neige, alors que j'étais perdu à l'INRA, coupé du monde civilisé, en proie aux affres d'une angoisse mortelle, m'a hébergé, m'évitant ainsi de périr par la morsure acérée du froid ou du loup ou de l'ours (c'est que c'est dangereux thésard comme métier, on risque souvent sa carcasse), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Raúl Zurita Milla (ça y est, j'ai commencé à lire les aventures de l'ingénieux Hidalgo), qui s'est cassé les dents et la tête sur la moquette (elles étaient pas assez grandes, trop ou pas assez blanches, tout ça), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Denis Hollecker, co-bureau d'un moment et stagiaire sur le susdit projet AP et qui a par conséquent fait de l'assimilation dans STICS (c'est que c'est pas donné à tout le monde) et côtoyé mes pantoufles, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Carlos Di Bellatchitchix[6] et sa famille Di Bellatchitchix pour sa personnalité toute personnelle et pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Emmanuelle Garrigues, la dame au chat, une transfuge du bâtiment sol, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements ; les stagiaires de l'été 2003 (vous savez, celui où qu'il a fait si chaud, ma bonne p'tite dame et avec ça, qu'est-ce que je vous mets, ben toujours pareil voyons, un bon bout de gras et du mou), Marion Sicard, encore une qu'on a mis à Matlab sans lui demander son avis, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Loïc Casagrande, le bohême, intermittent du spectacle et de la recherche (il le cherche aussi), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Christina Hau, de la même tribu que le susdit Clement Atzberger et la susdite Eva Pohl (c'est une grande famille l'université de Trier (Trêves), ils passent tous par Avignon), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Michaël Nowakowski, c'est celui qui a utilisé la méthode dite simplexe comme vous l'allez voir, c'est que je l'ai fait travailler, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements ; et parmi ceux que j'ai eu la chance de connaître un peu plus, et qui ont participé à la correction de ce document, Jérôme Demarty, malgré le grave différent qui nous a opposé sur un sujet fondamental et qui remonte au moins à la nuit des temps, ou du moins à la nuit de l'homme, à savoir la façon de préparer la soupe au fromage, qui n'est pas, comment certains oseraient l'insinuer, de la vulgaire soupe avec trois brins de fromage qui se battent en duel dedans (d'autant plus que tout le monde sait que pour se battre en duel il faut être deux, c'est mieux), lui défendant la version Aveyronnaise, moi la Tarnaise, deux partis inconciliables donc, initiateur du fameux saut de biche au dessus du (ou dans le) canal, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Cédric Bacour, grand spécialiste comme jamais on n'en vit à l'INRA de termes tels électro-pop/ électro-pop lénifiante/ powerpop/ rock/ pop-rock/ fusion-électro/ powerpop/ blues-rock-garage/ new wave/ rock-garage/ ragga-dub-indou/ pop-rock-électro/ soupe-pop/ folk-mou-du-genou/ rock énervé/ rock-électro dont les noms chantent la Provence ou peut-être pas vraiment en fait, et surtout du stoner (je croyais que ça s'écrivait stoneur ; cela dit, je reste persuadé que Bach est l'ancêtre du stoner), ainsi que de Matlab, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Sébastien Garrigues, mon compatriote, amateur de foisses (ils savent pas ce que c'est ici évidemment) et spécialiste reconnu de la tartiflette (un conseil, vaudrait peut-être mieux que tu cites pas trop ma thèse, ça pourrait faire vilain, et puis vérifie les codes des semi-remorques machin (et puis travaille pas trop tard)), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Franck Oro, qui m'a permis d'un peu mieux comprendre ce qui se passe en Côte-d'Ivoire, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Stéphanie Juin et ses boîtes en verre avec des bouts de canard dedans, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Kawtar Hdadi, une autre retrouvaille de Laon 2000 (ça marche au demeurant avec toutes les années ce jeu de mot à laonnois), et qui d'ailleurs y retourne, non mais je vous dis les déménageurs entre Laon et Avignon doivent y perdre leur Latin (tant qu'ils ne perdent pas les valises), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Nadia Rochdi, une retrouvaille d'Avignon 2000 quant à elle, sémillante Marocaine, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Mohammed Mumen, expert en géographie lybienne (il est Lybien, c'est normal, mais pas tant que ça après tout je me rappelle jamais le nom de la rivière entre la Bosse et la mer des Alpes) et en Islam (c'est bien la recherche, on voyage à moindre frais et en restant chez soi (à son travail) et en pantoufles s'il-vous-plaît), qui comprend tout en Français sauf mes âneries mais il ne perd rien, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Sophie Lebonvallet, encore une qui a fait Avignon-Laon et puis pouf, la Bolivie, allez savoir pourquoi (la pôôôôvre, devoir faire STICS quinoa), pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Matthew (Mafiou ; Mafsiou ; Matiou ; Masziou ; Massiou ; hum, pas bien fastoche à dire cette histoire) Pringle, dont j'ai susditement parlé en Anglais dans le texte siouplé, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements (même s'il aime pas le rugby, donc), Iñaki García de Cortazar y Atauri

[6] Allusion très peu sibylline aux écrits de M. Goscinny

y Quero y García y García (c'est fou le nombre de García dans ce pays) y Alcubilla y Hurtado de Mendoza y Lodares y Simon y Rojo y Calderon y Ansejo y Soupaloignon y Crouton y Dansonsurlepon y Davignon[7] (quelque chose me dit qu'il va dire en lisant ceci (si jamais il n'a pas autre chose à faire) : « Que cabron », ce qui doit être, entre nous soit dit, assez peu châtié niveau correction du langage hispanisant et dont je ne connais nullement l'exacte orthographe, n'étant pas un grossier personnage, surtout en Espagnol, comme me l'a assez expressément recommandé ma Maman) et, à ce propos, qu'il me soit permis de profiter de cette libre tribune ouverte à toutes les digressions, pour clamer encore une fois haut et fort ma totale et complète (ce qui revient un peu au même, j'entends bien, mais c'est un subterfuge de langage par le truchement duquel on insiste sur le sujet et que l'on peut nommer périssologie mais on n'est pas obligé non plus, on vit dans un pays libre après tout, sauf pour les pastèques dont les droits les plus élémentaires sont proprement bafoués) opposition au jeté de pastèque du haut (enfin, du milieu) du Pont du Gard, et je voudrai même aller plus loin, profitant de cette journée ensoleillée et retrouvant les accents d'un Jean Jaurès mort avant d'avoir eu le temps de le faire, embrasser la cause des pastèques du monde entier, qu'elles soient baies ou grains, entamer une lutte sans merci (de rien, vraiment, il n'y a pas de quoi) jusqu'à la mort s'il le faut pour faire valoir les droits de la pastèque, de toutes les pastèques sans distinction de quoi que ce soit ni que ce fut, ces pastèques qui ne sont pas des objets gestuels qu'on peut jeter du haut (ou du milieu, c'est pareil) du Pont du Gard comme ça et ce afin que jamais plus, ô grand jamais, une seule pastèque ne se sente forcée de se laisser lancer du haut du Pont du Gard (ou du milieu, donc, on s'en fout, donc) ; sans compter qu'il aurait pu y avoir quelqu'un dessous, je viens à peine de penser à ce détail, et qu'en jarre, en outre pardon, ça fait de la pollution nitrique, les pastèques dans le Gardon, c'est bien connu, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Marjorie Domergue, pour ses pastèques (dont une a été jeté du milieu du Pont de Gard d'une façon éhontée, comme narré susditement ci-dessus, *cf. supra*), sa bonne humeur enthousiaste et sa gentillesse et Cédric Bellenger, même s'il était un suppôt du privé (mais dans le bio, voyez-vous, ça rattrape, même si le bio, comme il est bien connu, ça file la … ; Ah bon ? On peut pas là ? Ça ne se fait pas ? Non ? Bon, tant pis, c'est dommage, je l'y aurai bien vu), qui a pris le relais de Marjorie dans le colportage matutinal de ma viande (ainsi que de mes os) vers l'univers enchanté (moi de même) de l'INRA, vu que voyez-vous, je voulais arriver plus tôt, eux qui furent mes voisins pendant quelques mois, pour leur porte toujours ouverte (même si ça fait des courants d'air), les apéros, les barbe-culs, sans compter les parties vespérales de jeux de plateau sur la table basse (vraiment trop basse) etc., etc., pour leur respect des fondamentaux, leur rigueur, leur discipline, leur engagement dans les regroupements, Clement Atzberger dont j'ai également susditement parlé ci-dessus (*cf. supra*) pour des raisons professionnelles en Allemand avec des fautes siouplé, lui aussi converti à la simulation plane des développement civilisatoires, pour son respect des fondamentaux, sa rigueur, sa discipline, son engagement dans les regroupements, Gaël Derive et Boris Ruelle (je les mets ensemble, c'est parce qu'ils se ressemblent tellement, mais n'allez pas jaser non plus), qui eux aussi me prirent en pitié et m'évitèrent ô combien de fois d'avoir à me commettre avec la plèbe dans les véhicules autobus publics en acheminant régulièrement et à peu de frais ma carcasse entière vers les douces rives du canal de l'INRA dans lequel il fait bon sauter à pieds joints, tous deux Grenoblois (c'est pour ça qu'ils se ressemblent) ayant fondé une colonie dans les remparts d'Avignon, colonie vouée à s'éteindre du reste (chacun en pensera ce qu'il voudra, mais c'est bien triste), autres amateurs de jeux de plateau (« je suis bien content pour toi si tu m'anéantis, tu sais, tu le mérites »), pour leur respect des fondamentaux, leur rigueur, leur discipline, leur engagement dans les regroupements (quoiqu'au basket, le gars Gaël il aime pas trop les engagements, hein ?), David Beal et Claire Lauvernet, un bien sympathique duo stéphanois, amateurs d'ail et d'oignon respectivement, de pions verts pour les deux mais de crème fraîche que pour une moitié, pour leur porte toujours ouverte (ou bleue ah ah) et les apéritifs aussi, *etc.*, *etc.*, sur lesquels il y aurait bien d'autres choses à dire mais on va pas couper trop d'arbres non plus, pour leur respect des fondamentaux, leur rigueur, leur discipline, leur engagement dans les regroupements (… n'allez pas croire, à moi aussi ça me gonfle cette petite litanie leitmotivesque) ; et ceux de la fin, juste entre-aperçus, finalement, Aurélie Girault, Paul Bernard, Philippe Rosselo, Vincent Rivalland, qui a aussi fait partie des gens qui ont corrigé ma thèse, si tôt arrivé, hop, au travail, c'est que ça rigole pas, et également plateauïste à ses heures, Kathie Pavageau qui aurait, à ce qu'il paraît, des talents celés en matière de narration d'histoire drôles (et pas des blagues Carambar apparemment) et puis bon, la toute nouvelle cuvée de stagiaires, mais, à l'heure à laquelle j'écris ces lignes, c'est-à-dire 22:00 précises en ce qui concerne celle-là, tiens, 22:01 maintenant, comme le temps passe, je ne sais pas forcément encore leurs noms de famille alors ça me va pas de citer des moitiés de noms, je n'aime pas faire les choses à moitié, je suis comme ça, faut pas chercher à comprendre. Je crois que le moment est propice (et quand je dis propice, ça ne veut pas dire que je suis pour l'humour pipi caca, hein, ça suffit à la fin) de remercier et de saluer comme il se doit (même s'il n'était pas à proprement parler un commensal régulier, mais ça lui fera plaisir d'être avec les jeunes) la star locale du chaud business : Bernard Seguin, qu'on ne présente plus. Je rappellerai superfétatoirement ses nombreuses interventions sur les ondes hertziennes et numériques à propos du réchauffement ou du refroidissement climatique (c'est selon l'humeur du moment)

[7] Allusion non moins peu sibylline aux ouvrages de M. Goscinny

occasionné par la chute des pingouins macchabées sur le dos ou sur le ventre, nous les avons tous suivies avec attention et intérêt (superfétatoirement n'existe pas, je sais, assez d'arguties superfétatoires justement, ne me détournez pas de mon sacré devoir, de mon devoir sacré pardon, et d'ailleurs c'est un mot indispensable, ce n'est tout de même pas de ma faute si l'Académie Française ne fait pas son devoir, l'exemple du laisser-aller vient du plus haut, regardez, on ne sait toujours pas si on doit dire paramétrisation, paramétrage, calibration, calibrage, calibrationnement, étalonnage, étalonnationnement, bidouillage, ou même encore paramètres constants, paramètres spatialement variables, paramètres temporellement variables, variables temporellement constantes, variables spatialement constantes, variables spatio-temporellement constantes mais paramétrisables temporellement et spatialement, ou encore parammouths et variabmouths ; si vous cherchiez (j'ai bien dit : cherchiez) ce mot (superfétatoirement) dans un dictionnaire, vous risqueriez de vous fourvoyer entre superfétation et superficialité, voire même, si vous poussiez votre rêverie philonyme (encore un qui n'existe pas, mais qui est-ce qui m'a foutu cette langue ?) quelque peu plus loin (ou moins loin tant il est vrai que chaque page d'un dictionnaire tournée avant celle qui contient le mot recherché risque de vous attirer dans ses propres méandres et vous faire oublier icelui); vous risqueriez de tomber sur étrivière, lycoperdon, voire même valétudinaire ou cacochyme, ces derniers n'étant naturellement pas une allusion à Bernard Seguin, allons donc, voyons, voyons, ne m'attribuez point mots qui ne sont miens, tss, tss ; et là, bon, je consens à refermer cette parenthèse que d'aucuns jugeront trop longue), mais pas sa jovialité non entamée par ses hautes fonctions, vertu rare, ce dont je ne saurai que lui être redevable (j'arrive un peu à épuisement de la liste). Je me permettrai cependant de lui recommander de parler moins fort, c'est pas facile quand on travaille, vous ne vous rendez pas compte (et vous pouvez évidemment toujours vous gratter pour que je me mette à vous tutoyer, et d'abord ma Maman ne serait pas plus d'accord sur ce sujet que sur celui qui consisterait à ce que je m'aventurasse à tenir je ne sais quels propos outrecuidants et castillans, du genre que cabron, que je saurais même pas ce que ça veut dire et elle non plus d'ailleurs je suppose (ou alors elle me cache des choses)). En arrivant vers la fin de cette litanie égrenée avec une indicible émotion, j'exprimerai également ma reconnaissance envers Bénédicte Quilot qui n'a jamais refusé de répondre, étant de peu mon aînée, à des questions saugrenues du genre : « Et comment fait-on pour s'inscrire sur le site d'ABIES et est-ce vraiment bien nécessaire » (pas aussi bête les questions mais presque ; aussi, on n'a pas idée de me traumatiser avec des tracasseries pareilles) et Norbert Turion (encore un qui vous réconcilierait un écolo borné avec les chasseurs) et, tant qu'on en est à citer de vieilles connaissances, Mathilde Cohen et Sylvain Chabet-Ferret, un de ma tribu mais d'origine picarde et Christophe Raoul (« Et vous, vous en avez vu, des électrons ? Eh bé non, et pourtant, vous êtes bien obligés d'y croire, moi non plus, j'en ai pas vu, et j'y crois. Et le petit bonhomme, là, qu'est-ce qu'y fait ? Eh bé, il les fait passer, aux électrons, dans la petite boîte. / T'es sûr de toi, là ? … Eh oui, c'était ça, fallait pas effacer, allez, va, continue, tu l'avais un peu mis au pif, non ? »), ainsi que Mélanie Tauber et Roberto Garcia y quoi déjà ? ah oui, y Alcubilla (y a pas de « y » d'ailleurs je crois, je t'y comprends rien à leurs salades espagnoles), chez lesquels j'ai trouvé un pied à terre à Paris lors de mes déplacements professionnels (oui, mesdames, mesdemoiselles, messieurs, j'ai bien dit professionnels) en cette dive métropole, notamment à l'occasion de ma soutenance, et l'Empereur du Japon (ça faisait longtemps) naturellement (c'est un peu un leitmotiv récurrent (avec un e et deux r, pas un a et un r évidemment) celui-là, comme les ours ; l'Empereur du Japon un leitmotiv ? Voyons, jeune homme, un peu de respect) qui, non content d'avoir quitté son Empire du Soleil Levant (levant quoi ?) aux doigts de rose pour la première fois pour aller visiter Albi[8] s'est écrié, au pied du donjon de la basilique-forteresse (non, pas : « Du haut de ces pyramides, quarante siècles vous contemplent », c'est pas de lui ça, mais d'un autre empereur qui ne l'était pas encore) : « Ouh là ! Bigre[9] ! J'aimerai pas choir de si haut dites donc » (peut-être pas si intello que ça finalement l'Empereur du Japon), montrant ainsi, par cette allusion sibylline à la mort du grand Henri, sa parfaite maîtrise de la culture norroise (ah oui, tiens, j'avais pas vu tout ça dans un premier temps, chapeau l'Empereur du Japon ; oui, les lectures des billevesées utopiennes, hautaines, compassées, redondantes et pontifiantes, véritables monuments à la pensé unique, ça attaque un peu même si ça fait marrer quand on est de bonne humeur).

Et enfin, je remercierai tous les membres de ma famille parce que hein, bon, quand même, tout de même, évidemment, voyez-vous quand même hein, et oui tant que ça évidemment alors là, hein (je viens de relire cette dernière phrase (parce qu'il ne faut pas croire, mais je me relis quand même parfois même si je n'en donne pas l'impression comme ça) et je me demande si le sens profond sous-jacent à mon propos est bien intelligible en l'absence de l'intonation approprié permettant d'infléchir le ton de la voix avec un x bien sûr et créant la signification, la portée oserai-je même dire, du verbe humain, transcendant les mots) mais je ne vais peut-être pas tous les citer non plus parce que là on ne s'en sort plus (tiens, l'autre jour, rien que du côté de mon père, on était 96 sans forcer), alors au moins ma Maman, mon Papa, mon frère et mes sœurs (trois dans l'Aube)

---

[8] Fallait quand même bien que j'en parlasse

[9] En Français dans le texte

et évidemment les rapportés comme on dit avec une poésie qui nous honore, ainsi que les petits bouts de chou de neveux et nièces qui sont bien rigolos, même s'ils nous poussent inexorablement.

*
***
*

Doit-on remarquer que le texte qui va suivre n'est pas écrit en vers et l'est donc, propriété corollaire, en prose. Non. Mais comme il se trouvera toujours des esprits chagrins pour le faire, autant que je le fasse moi-même, dans une sorte de *mea culpa*, on n'est jamais aussi bien servi que par soi-même dit l'adage populaire, surtout en matière de critique, ça permet de s'arrêter au bon endroit et au bon moment, ça c'est moi qui le rajoute, assez marqué par l'agriculture de précision comme vous pouvez le deviner à ce leitmotiv revenant comme une rengaine. Il faut toujours écouter le peuple, surtout ses grands-mères, en général férues d'adage. Donc, je l'avoue sans vergogne (et sans vers tout court donc, ni plus de cigogne) : j'ai écrit le document suivant en prose, parce que j'ai eu la flemme de l'écrire en vers et que de toute manière je n'ai jamais su écrire en vers. Si jamais on devait me reprocher de n'avoir pas écrit ce qui suit en vers, ou en Latin, je ne prétendrai jamais tenter – ce serait ô combien vain – de m'inscrire en faux contre une si pertinente, si judicieuse, bref, si juste remarque. Je n'essaierai pas même de m'en défendre ou de trouver quelques fallacieuses excuses. Mais, pour lever cette critique méritée, j'utiliserai le fameux et quelque peu éculé prétexte que nous serons toujours à temps de faire des vers une fois mort, ce qui fait que nous pouvons très bien nous dispenser d'en faire de notre vivant. Et, en cette heure de la journée où je retrouve du poil de l'ours ; tiens, non, de la chèvre (oui, déçu par ces sympathiques plantigrades qui n'étaient décidément pas à la hauteur de mes espérances (cette sibylline allusion à l'ours est, je m'en excuse, quelque peu private joke (et celui qui s'avise de prononcer prail vit'h djauk, je lui envoie un ours et une chèvre lui dire ce que je pense, c'est-à-dire pas grand chose, vu que ça me fatigue de penser)), j'ai maintenant jeté mon dévolu philanthrope sur ces non moins sympathiques artiodactyles fixant sur le monde leur pupille oblongue et ruminant d'antiques apophtegmes dans leur non moins sympathique barbe), ayant retrouvé du poil de la bête donc, je me hasarderai à citer le propos que Cervantès tint en guise de phrase liminaire de son Don Quichotte qui allait devenir un des chefs-d'œuvre de la littérature mondiale (ce que disant, je ne m'engage guère, eu égard au fait que je ne l'ai pas encore lu de façon exhaustive, donc, je ne fais que bêtement répéter comme toujours et comme Gonzague ce que j'ai dû lire, ou entendre, ou imaginer, ou rêver, mais il y a quand même des raisons d'imaginer que si Cervantès figure sur les pièces de 50 cc d'Euros espagnols, c'est qu'il a dû être un fameux personnage et pour autre chose que le tricot ou le crochet (dont les adeptes ont finalement assez peu laissé de traces de leur passage sur terre au cours de l'Histoire) mais évidemment un peu moins fameux que celui qui siège sur celles de 1 et 2 €, et dont le nom m'échappe d'ailleurs, comme quoi l'argent ne fait pas la gloire ; la Didi peut-être ? Non, c'était en Angleterre, ça), l'énoncé que voici que voilà, ça vient, ça vient : « Lecteur inoccupé, tu me croiras bien, sans exiger de serment, si je te dis que je voudrais que ce livre, comme enfant de mon intelligence, fût le plus beau, le plus élégant et le plus spirituel qui se puisse imaginer ; mais hélas ! je n'ai pu contrevenir aux lois de la nature, qui veut que chaque être engendre son semblable. Ainsi, que pouvait engendrer un esprit stérile et sec comme le mien, sinon l'histoire d'un fil sec, maigre, rabougri, fantasque, plein de pensées étranges (…) ? ». Donc, si Cervantès, de glorieuse mémoire, ayant été ce qu'il a été, a cru bon de tenir semblables propos à la lisière de l'ouvrage qui tressa ses lauriers, combien plus, nous autres, nous devons-nous de faire amende honorable et de prier à genoux le lecteur qui est, lui, à l'inverse de celui auquel Cervantès s'adressait et qui s'apprêtait à lire son ouvrage dans le but de se distraire, ô combien occupé, un peu comme les toilettes lors d'épidémies aiguës de diarrhées, ayant force autres travaux à abattre et ô combien plus profitables que de lire ce qui suit, combien, donc, nous devons-nous de le prier de bien vouloir excuser les mille et mille imperfections et vices dont sont souillés nos pauvres manuscrits relatant le fruit de nos indigentes élucubrations tant vespérales que matutinales. Mais, si le texte qui suit est ce qu'il est, c'est pas ma faute. et non, c'est la faute à pas de chance ou à la fatalité voyez-vous, ou comme vous voudrez le nommer, et laissez les ossements et la mémoire de Voltaire hors du coup, comme je vais vous le démontrer sans plus attendre avec force arguments (à l'inverse de personnes qui ont la désagréable marotte d'assener à l'emporte-pièce leurs opinions à grand renfort d'effets de manche et avec comme seuls étaiements de leur propos des moues dubitatives associées à des haussements d'épaules) irréfragables, ou du moins vais-je le tenter avec l'entrain naïf et l'enthousiasme propre à mon jeune âge.

Ne vous est-il jamais venu à l'esprit, non sans quelques angoisses sans doute, que l'instant où vous commencez à écrire un texte, les conditions dans lesquelles vous l'écrivez, en influençant vos dispositions à le faire, conditionnent largement la qualité de votre écrit à venir ? Ne vous est-il jamais apparu avec la plus fulgurante des évidences que, si vous aviez commencé à écrire une seconde plus tôt ou 3h 27'48'' plus tard, votre texte eût été de bien meilleure facture ? Parce qu'alors, vos pensées se fussent peut-être mieux enchaînées, organisées et votre vision d'ensemble du problème à traiter eût été plus claire ? Parce qu'alors, votre période de

réflexion et de composition eût comporté une de ces rares secondes de fugace clairvoyance au cours desquelles les problèmes apparaissent, limpides, dans tout leur ensemble ? De façon corollaire, on peut aussi penser qu'à d'autres moments, le résultat du travail d'écriture eût été encore plus mauvais. Donc le moment où on écrit un texte conditionne sa qualité, de façon incontrôlable. Pour aller plus loin, un texte quelque peu long doit nécessairement être écrit par étapes, ce qui rend difficile de satisfaire l'impérieuse nécessité d'en faire un ensemble uni, homogène et cohérent. Si les textes pouvaient être écrits dans leur intégralité au cours de ces fameuses secondes, ces secondes qui ne se commandent pas, ils pourraient peut-être naviguer dans les eaux de l'indicible perfection de l'entendement. Mais ces textes parfaits, existent-ils seulement ? Oui (ah, c'est bref, comme réponse). Je m'en explique (ah, bon). Les textes parfaits – c'est-à-dire les meilleurs qui puissent s'appliquer à décrire ou traiter un point particulier – existent car tous les textes existent (allons bon, vous piquez ma curiosité, ce que disant), les bons, les moins bons, les mauvais, les pleins de fautes, et ce, dans toutes les langues qui furent, qui sont ou qui seront. Les textes sont en effet des combinaisons de caractères, donc des nombres (les Grecs anciens n'utilisaient-ils pas toutes les lettres de leur alphabet en guise de chiffre ?) : il suffit d'assigner à chaque caractère typographique une valeur numérique et compter dans la base appropriée. Les textes sont des nombres, or, tous les nombres existent, donc tous les textes existent (si c'est pas du syllogisme à l'ancienne et aux petits oignons (rajoutez thym, romarin, coriandre et cumin), ça, je t'y comprends plus rien). Tous les textes existent, insisterai-je quitte à me répéter, ceux qui ont déjà été écrits, ceux qui le seront, plus une foultitude qui n'ont aucun sens, à nos yeux bien sûr (parce qu'il y a aussi les textes des Martiens dans le lot). Il existe même, par exemple, des versions de *Notre-Dame de Paris* avec des fautes partout, une signée de mon nom (alors même que mis à part les initiales je n'ai pas l'honneur de partager quoi que ce soit avec son véritable auteur), du vôtre, des textes de recettes de cuisine succulentes et qui ne seront peut-être jamais inventées, quel dommage, et le texte parfait pour cette thèse. Pour trouver les textes parfaits, il suffirait donc de vérifier un à un les textes qui existent, il suffit de savoir compter (un peu plus loin que jusqu'à mille contrairement à mes nièces), de choisir un à un les textes, d'être assez patients et d'avoir du temps devant soit, car ils se trouvent en nombre infini, ce qui complique, j'en conviens, singulièrement la chose pour des êtres qui ont nos limites. Bref, je n'étais pas peu fier de cette trouvaille (que je n'ai pas essayée de mettre en œuvre), lorsque j'avisai que Swift en avait formulé une proche dans ses *Voyages de Gulliver* : il y met en scène un savant qui prétendait faire avancer la science en mettant, grâce à une machine (déjà, avant l'invention de l'ordinateur, quel précurseur), bout à bout, selon un processus aléatoire, des mots pour former des bouts de phrases, dont des disciples (déjà, avant l'invention des stagiaires, quel précurseur) extrayaient ceux d'entre eux qui étaient intelligibles. Swift adoptait donc là une approche stochastique si je puis dire, là où je propose une approche d'exploration systématique de l'espace des textes existants, mais l'idée est bel et bien la même, et il est probable que des foules de gens l'ont déjà eue. Ce qui prouve bien qu'il est difficile de penser à quelque chose d'original. L'Ecclésiaste (1, 9-10) l'avait bien vu à son époque reculée : « Ce qui fut, cela sera, / ce qui s'est fait se refera, / et il n'y a rien de nouveau sous le soleil ! / Qu'il y ait quelque chose dont on dise : « Tiens, voilà du nouveau ! », cela fut dans les siècles qui nous ont précédés. », et plus ça va, moins ça va, bien sûr, puisque chaque époque a apporté sa petite pierre à l'édifice. Donc, tout ça pour dire que si je suis passé à côté du texte parfait, ou, pour viser un objectif plus modeste, d'un bon texte, c'est parce que, vraiment, il y a trop de nombres, et rajoutez à cela qu'on risque toujours de perdre le compte, et l'infini, ça me fait un peu peur, surtout quand j'y pense le soir au fond de mon lit comme quand j'étais un tendre enfant entre les cils duquel se dessinait un ballet de points rouges, verts, jaunes et bleus en farandoles mouvantes et en bancs virevoltants dans l'obscurité tombée, qu'y puis-je ?

Un autre argument pour expliquer la qualité douteuse de ce qui suit, est que, d'une certaine façon, nous sommes aussi des nombres, par exemple ceux qui définissent la position de chacun de nos atomes par rapport au référentiel que vous voudrez, choisissez, je sais jamais choisir, et puis c'est égal puisque tout est relatif paraît-il alors pourquoi choisir justement. Comme nous bougeons (mais si peu au regard de l'Univers, finalement), ces nombres bougent aussi au cours du temps, mais peu importe, il suffit de mettre bout à bout les coordonnées de chacun de nos atomes à chaque instant de notre vie, et on ne sera défini que par un seul nombre. Et, ce poulet, ce faisant, pardon, on pourrait connaître, par le truchement de cet expédient, notre avenir, si on savait à quel nombre on est ainsi lié. D'ailleurs, dans tous les textes qui existent, qui sont aussi des nombres donc et donc peut-être des personnes, il y a aussi celui qui raconte notre vie, mais ce ne sont pas forcément les plus intéressants surtout en ce qui me concerne. Bref. Or, il est bien connu qu'aux chiffres, on leur fait dire ce que l'on veut, donc mon chiffre ne m'a jamais fait dire que ce qu'on lui a dit de me faire dire, donc c'est pas ma faute, mais encore faudrait-il savoir ce qu'est ce on. Un peu dans le même ordre d'idées (si tant est que nous ayons à faire là à des idées), on peut ressortir des tiroirs les vieux concepts d'animaux-machines et d'hommes-machines (homme au sens d'ανθροπος et pas seulement d'ανδρος), qui stipulent que tous nos gestes sont le résultat et le fait d'une mécanique bien réglée et bien déterministe. Si un animal se déplace, c'est parce que ses muscles bougent, s'ils bougent, c'est parce qu'un influx nerveux est venu déclencher la cascade de réactions qui sont nécessaires à ce mouvement, si cet influx nerveux leur est parvenu, c'est parce que son système nerveux central l'a envoyé en réponse à un stimulus, stimulus qui a été traité par ce même système nerveux selon des mécanismes qui sont régis par l'information génétique de ce susdit (ça faisait longtemps, c'est juste pour qu'on ne puisse pas croire

que soudain je suis sérieux) organisme. Il en va de même pour l'homme (ανθροπος, donc), il n'y a pas de raison, c'est un animal comme les autres, il a un peu pris la grosse tête, c'est tout. On peut poursuivre ce raisonnement en allant vers ce que d'aucuns ont pris l'habitude de nommer l'infiniment petit, car tout ceci est fractal voyez-vous, comme une thèse. Nous sommes de la matière et la matière est, du moins dans la limite de mes pauvres connaissances, « composée » d'ondes, dont les interactions et le comportement sont dictés selon des lois physiques. Nous serions donc nous-mêmes mués par les lois austères et irréfragables de la répulsion des ondes de même type et par l'attirance des ondes de types opposés. C'est uniquement parce que les ondes positives et négatives s'attirent que je viens de me gratter le nez (l'extérieur du pavillon, je vous rassure, j'ai un peu de tenue, allons donc, malgré mes ondes rustres). Tout l'univers ne serait qu'ondes grouillantes positives et négatives s'attirant et se repoussant selon un principe binaire, booléen, si alors si et puis c'est tout. Et les ondes n'ont aucune dimension, aucune épaisseur, aucune consistance, aucun galbe, aucune chaleur, elles grouillent et, disent certains, à condition d'y mettre l'énergie qu'on tirerait de je ne sais où, d'une sorte d'anti-gros-boum, ou alors par la seule force de l'esprit qui n'est jamais lui aussi que grouillement d'ondes aussi peu frivoles les unes que les autres, on pourrait, disent-ils, ces gens-là qui ont scrupuleusement étudié le problème à grand renfort d'élucubrations de leurs ondes et de gratouillements d'ondes avec d'autres ondes, ramener la taille de l'univers à la taille d'un verre d'eau, pourquoi un verre d'eau, quel hasard que la taille de l'univers écrabouillé correspondît justement à la taille de cet objet, lui même paquet d'ondes grouillantes, et ce depuis toute éternité, bien avant l'invention du verre d'eau et *a fortiori* du verre à dents, pourquoi pas un dé à coudre, et puis il y a verre d'eau et verre d'eau, ils me font bien rire avec leurs recettes où ils mesurent les volumes avec des verres d'eau, il y en a de toutes les tailles, comment voulez-vous respecter les fondamentaux de la rigueur culinaire avec ça, et quand on boit un verre d'eau on boit l'univers, quand on chausse un dé à coudre, on écrase un peu plus un univers déjà mal en point, et en rajoutant encore un peu d'énergie, l'énergie du désespoir sans doute, parce que quand l'univers est réduit à la taille d'un dé à coudre, l'espoir doit bien être réduit à peau de chagrin, en rajoutant un peu d'énergie, puisque les ondes n'ont d'autre consistance que leur manie de se repousser, de s'attirer, de grouiller et de girouetter, on pourrait le réduire à néant, que deviendrait alors le dé à coudre personne n'en parle étonnamment, l'univers est fait de dés à coudre, on mesure l'univers à l'aune du dé à coudre, nouveau système métrique. Ainsi donc, si je viens d'écrire ce que je viens d'écrire et qui n'est pas très spirituel j'en conviens, c'est uniquement parce que les ondes qui me font l'honneur de constituer ma viande et mes eaux si j'ose ce jeu de mot laid, sont mues par ces mêmes lois (et la loi, c'est la loi, papiers du véhicule sivouplé) de répulsion et d'attirance, et, à force de s'attirer et de se repousser selon ces lois, elles en sont arrivé à faire en sorte que le paquet d'ondes qui me sert de stylo a laissé traîner sur un autre paquet d'onde peu large des lignes d'autres paquets d'ondes. Tout ceci est de la physique ; revue et visitée j'en conviens encore. Aussi bien, et c'est là que je voulais en venir, si le texte qui suit n'est pas bon, vous sentez bien que je n'y suis pour rien, car ce serait vouloir me faire endosser un drôle de chapeau (alors qu'un chapeau se chausse), la responsabilité de la législation universelle, et ce serait quand même y aller un peu fort, je (enfin, mes ondes) ne mérite pas tant d'honneur, j'étais même pas né à l'époque. Autrement dit, avec un bon modèle (un modèle parfait) et en connaissant parfaitement les conditions initiales régnant aux débuts de l'univers, on aurait pu prévoir que j'allai écrire, grand dadais que je suis, des trucs du genre : « On étudierait ensuite le comportement de la différence entre les valeurs de critère pris pour les doses variables et pour la dose unique. » C'était à ce moment-là qu'il fallait agir, maintenant, la mécanique est lancée, comme dans les tragédies. Mais il faut croire qu'ils avaient d'autres chats à fouetter à l'époque.

Mais, las, tout ceci est affreux. Allons faire des crêpes car, aussi vrai que nous sommes de la pourriture en devenir, les crêpes c'est bon, surtout avec du sirop d'érable. « Après mon septième apéritif, je pensais qu'il n'y a ni réel, ni irréel, ni vérité, ni mensonge. (...) C'est bien malin de philosopher sans avoir appris à philosopher et après sept apéritifs. » C'est de Ionesco et de circonstance, non point tant au sujet des apéritifs (j'en suis aux crêpes) que pour le fait de ne pas avoir appris à philosopher. Trêves (pas Trier) de ces visions sombres et sorties des pires développements eschatologiques et allons vers de plus riantes contrées. Et que ceux qui s'aviseraient de tirer argument de l'emploi de ce mot (eschatologique) pour alléguer que je suis miction fèces s'en remettent à leur dictionnaire La Rousse ou Roberte. Et qu'on veuille bien ne me point porter ni grief ni rigueur de ces légères fautes d'orthographe innocentes, franchement je n'ai pas le cœur à ce genre d'interrogations, et en plus on s'en fout, eu égard au fait que la première règle de l'orthographe est que les noms propres n'ont pas d'orthographe, surtout les noms étrangers qui n'en font vraiment qu'à leur tête. Ensuite, c'est tout parce que l'orthographe n'est jamais qu'une affaire d'usage. Pour la grammaire, c'est plus délicat : en effet, il ne faut pas jeter la grammaire dans les orties, parce qu'après elle n'est pas contente et grammaire pas contente, grammaire foutue. Je conçois que pour la plupart des gens ce point essentiel soit assez secondaire, à la limite du tertiaire, c'est le crétacé je crois, c'est assez en effet dit le dauphin, je fais le gros dos et me cache à l'eau, plouf, mais que voulez-vous ? Les gens sont comme ça de nos jours. Ils se moquent de tout. Et ça date pas d'hier d'ailleurs. Déjà du temps de ma mère, on tenait des propos du genre : « Ah les jeunes de nos jours c'est plus ce que c'était et avec ça y a plus de saison, cette jeunesse qui monte, quelle descente, un peu plus de boudin je vous prie mon cher monsieur ». Ce qui fait qu'on peut fort logiquement en déduire que les jeunes d'hier, qui sont les vieux d'aujourd'hui ou d'après-demain, c'était déjà plus ce que c'était, alors les jeunes d'aujourd'hui qui sont les

vieux de dans quelques temps, c'est de mal en pis, et par conséquent les vieux de nos jours c'est plus ce que c'était non plus et alors les vieux de dans deux ans, je veux même pas y penser, surtout s'il se met à faire des canicules tous les étés, non c'est vraiment plus ce que c'était. Mais ces réflexions risquent de fort peu nous importer au soir de nos vies respectives, nous apparaissant alors dans tout leur appareil de futilité. Et oui, mes bons amis, le temps passe vite, témoin mon teint buriné du fait de l'exposition réitérée aux traits dardés par l'Apollon tant picard que provençal (c'est pas tout à fait de moi ça non plus je crois) qui a foutu le camp et laissé la place sur ma peau d'albâtre au teint éburnéen d'un cénobite, eu égard au caractère hiémal de la saison (moi aussi hiémal à la tête) et à la rédaction fort occupante du texte qui va suivre. Non, monsieur ou madame, je n'ai pas bu le Bateau Ivre de Georges Rimbaud. Je dis ça soudain au cas où un fâcheux (et il s'en trouve toujours dans le nombre) s'aviserait, allez savoir pourquoi, de trouver un air de ressemblance quelconque entre cet affligeant soliloque et le Bateau Ivre de Raimond Rimbaud et de me reprocher, perfide comme tout fâcheux, un plagiat improbable vu que je me décarcasse à vous l'expliquer sur tous les tons, je n'ai pas bu le Bateau Ivre. Du reste, il est très peu probable (les revoilà, ces probabilités, comme quoi elles sont comme les Erinyes et le vengeur masqué, on n'y peut échapper) qu'il y ait quelque chose de commun entre le susdit présent radotage et le susdit poème. Mais vu que je l'ai pas lu, je n'en sais fichtre rien, alors au cas où, sait-on jamais voyez-vous, je prends mes précautions, on n'est jamais assez prudent avec les fâcheux. Tiens, j'ai écrit par deux fois bu en lieu et place de lu. Ce n'est pas grave, je garderai tout de cette crise logorrhéique, la logorrhée, comme chacun sait, étant au discours ce que la diarrhée est au transit intestinal. Donc, disais-je, je n'ai pas lu cette œuvre majeure, que voulez-vous, il en va ainsi de ces choses comme d'autres, mais c'est pure négligence de ma part, car il paraît que c'est trodlabal, comme disent les jeunes. Ils sont cnons ces jeunes. C'est pas ma faute, j'ai horreur des jeunes. Franchement, je vous le dis comme je le pense, tout net, comme ça, là, à brûle-pourpoint, de façon impromptue et tout à fait entre nous, je ne conçois même pas que de nos jours, en plein 21ème siècle, au troisième millénaire, bref, à l'âge d'or du nucléaire et du réchauffement climatique, il se trouve encore des gens pour avoir l'idée saugrenue d'être jeunes. On dirait que les gens ne tirent vraiment pas de leçons des riches enseignements de l'histoire. Et après, tous ces jeunes cnons, que vont-ils faire ? Que vont-ils devenir ? Hein ? Je vous laisse deviner : des vieux cnons et oui, bien sûr, vous avez trouvé monsieur, toutes mes félicitations, fallait quand même le faire. Ah, ça, ils l'auront bien cherché, après tout, car le meilleur moyen d'être cnon étant vieux, c'est de l'être étant jeune. Comme dirait Arthur Brassens, non, ce doit être Georges lui en fait, bref, ne chipotons pas superfétatoirement sur ce genre de détail (superfétatoirement n'existe toujours pas : mais que font donc ces fainéants de l'Académie Française ?), comme disait, à peu de choses près, Georges (René ?) Brassens : « Quand on est cnon, on est cnon ». D'ailleurs, comme ça, juste au passage, chapeau Georges (Robert ?) Brassens, vraiment, on se demande où il allait chercher tout ça, c'est un peu comme pour William (Averell ?) Prout, j'aurai jamais trouvé ça tout seul. Tenez, par exemple, moi qui suis encore jeune, n'est-ce pas, si, quand même, un peu, je vous assure, un peu aigri précoce peut-être mais jeune quand même, c'est pas pour me vanter, mais je viens d'en dire pas mal, là, de cnonneries. Preuve *exempli gratia* (et un peu *ab absurdo* aussi au passage) que les jeunes sont des cnons. Ah, mais me direz-vous, chipoteurs que vous êtes, c'est un syllogisme ça ou je sais pas quoi, mais une généralisation abusive, ou induction peut-être ? Bref un truc dans ce genre. Certes, mais que voulez-vous que j'y fasse, et franchement, j'ai pas envie de peaufiner la démonstration, c'est assez évident comme ça en pichet ; en outre, pardon. Bon, et puis, si je suis un jeune cnon, qu'y puis-je ? Je ne vais tout de même pas me transformer en vieux cnon précoce rien que pour vos beaux yeux, non ? Bon. Allons, quand même, reprenez-vous, que diable.

***

Là encore, j'avoue que je sèche un peu sur la transition, c'est pour ça que je mets des étoiles.

***

Xénophon, dont on peut légitimement se demander pourquoi j'en parle ici (mais cela répond à un projet trop contraignant et de ce fait abandonné), a dit : « Les actes ne sont-ils pas des témoignages plus dignes de foi que les paroles ? », mais encore : « Il n'est condiment que l'appétit » (l'appétit vient en mangeant, c'est d'un autre), ou encore : « Interroger, c'est enseigner » (oui, c'est très facile de trouver des citations pour faire le cuistre), mais on s'en fout et a, par ailleurs, conduit les 10000 (10000 au début ou à la fin ? Nous partîmes 10000 et arrivâmes 500 ; non, c'est pas eux, ça) au terme d'un long périple dont ils ont salué la (presque) fin par un cri du cœur : « Θαλασσα ! Θαλασσα ! », parce qu'ils avaient déjà la télévision et que la troisième chaîne diffusait déjà l'émission du même nom le vendredi soir, mais ce terme servait plus communément à l'époque à dénommer Thétis, personnification de la mer (il y avait des façons plus simples de dire ça), dans l'antique vocable grec ancien (il y a donc fort longtemps). C'est un peu ce cri du cœur qui s'exprime ici, à la différence près que mon odyssée a duré trois ans et non 1.5 comme ce fut le cas pour les 10000. Certes, à leur décharge et pour justifier l'œuvre de Xénophon (que je n'ai évidemment pas lue), je me dois de rajouter que je ne peux pas prétendre avoir eu faim et soif au cours de ces trois années de folles méditations et autres ratiocinations et vaticinations

bucoliques, et les autochtones m'ont à peu près laissé tranquille, du moins pour ce qui est de me darder de coups de pilum (ce qui constitue un contresens puisque tout aussi bien les pila (pili ?) armaient l'armée romaine et non les armées des autochtones qui harcelèrent les 10000. On me pardonnera donc ces égarements.

***

Pour en terminer ou presque avec ce salmigondis indigeste d'élucubrations diverses, je tiens à préciser à l'intention des esprits chagrins s'il s'en trouvait jamais parmi les lecteurs (ce que personnellement je ne puis souffrir de me résigner à croire), que j'ai rédigé ce bref soliloque en dehors des heures de service et de repas car je n'en saute jamais, c'est contre nature. Enfin, avant de nous acheminer vers notre destin funeste (pourquoi funeste ? Parce ce que ça fait joli, ça fait homérique, oui monsieur, madame, parfaitement, voilà), en lieu et place de la ritournelle afférente à la perpétuation de l'espèce qui clôt traditionnellement les remerciements de thèse, voici, et comme il sied à tout texte de haute volée scientifique digne de ce nom et qui se respecte, le résumé en Anglais, autrement dit « summary » ou « abstract », dans lequel je m'en vais tracer à grands traits les points essentiels de ce qui précède, et ce à l'attention du lecteur pressé, et Dieu sait s'il y en a de nos jours (on n'a plus le temps de vivre madame, eh oui, je vous remets du mou ?). Mais, afin de ne pas choquer les farouches partisans de la langue de Molière et de Durant qui auraient bien préféré (les farouches en question) que leur langue et non cette dernière s'imposât au monde et qui du coup crient au scandale, le susdit abstract se voit précédé d'un résumé.

Résumé :

L'auteur a remercié à grandes enjambées les personnes qui l'ont aidé pour son travail, et notamment son altesse impériale l'Empereur du Japon, qui doit bien se demander ce qu'il fait là, mais on peut douter qu'il l'apprenne jamais un jour. Evidemment, comme il est latin, ça a pris plus de place que le strict nécessaire. Enfin, que voulez-vous, c'est ainsi. Ensuite, assez fourbement, il a tenté de persuader le lecteur (qui, malin, ne s'y sera pas laissé prendre) que si ce qui suit n'est pas à la hauteur de ses espérances, ce n'est pas de sa faute à lui, mais de celles des ondes qui grouillent et que par ailleurs l'Ecclésiaste l'avait bien dit que ce ne serait pas bien. Au passage, il a fait preuve d'une évidente cuistrerie dont il ne se cachait même pas, saupoudrant son texte de citations visiblement incomprises, tant leur emploi était manifestement hors de propos, un peu comme pour l'Empereur du Japon dont, à la réflexion, on ne voit toujours pas pourquoi il serait venu ramener ses guêtres (si tant est qu'il en porte) ici. Enfin, il s'est lancé dans un réquisitoire tout à fait inutile et vain contre la jeunesse d'hier et d'aujourd'hui et l'Académie Française, qui n'a toujours pas statué à l'heure h sur l'invention d'un mot pour désigner les paramètres qui varient. Enfin, il termine en saluant Xénophon comme il se doit, parce que c'est un jolly good fellow (Xénophon).

Abstract: (comme c'est un peu compliqué à traduire, j'ai confié la tâche à un traducteur automatique)

The author thanked with great strides the people who helped her for her (hmm, here I must correct myself and precise that it must be "his" understood) work, and in particular his imperial highness the Emperor of Japan, who must well wonder what he makes there, but one can doubt that he will never learn it a day. Obviously, as he is Latin, that took more place than the bare essential. Lastly, what do you want (as would said the President), it is thus. Then, rather perfidiously, he tried to persuade the reader (who, malignant, will not have let himself there take) that if what follows is not with the height of his/her hopes, it is not his fault with him, but of those of the waves which swarm and that in addition Qoheleth had said it well that it would not be well. In the passing, he showed an obvious priggish pedantry of which he did not even hide, powdering his text with obviously misunderstood quotations, so much their employment was obviously out of matter, a little as for the Emperor of Japan which, with the reflexion, we still do not see why he would have come to bring back his gaiters (if as well is as he carries gaiters) here. Lastly, he launched out in a completely useless and vain indictment against the youth of yesterday and today and the French Academy, which still did not rule at the hour H on the invention of a word to name the parameters which vary. Lastly, he finishes by greeting Xenopho as he should be, because he is a joyeux bon camarade (Xenopho).

« A chacun son pet sent bon. »
Erasme, comme quoi, l'exemple vient d'en haut
(cela dit, si je puis me permettre, je ne suis pas tout à fait d'accord sur ce point avec ce grand homme)

Table des matières

Premier Chapitre : Contexte de l'étude, objectifs et démarche p.23

I.1 Introduction p. 25

I.2 Fertilisation azotée et agriculture de précision p. 26

I.3 Présentation de l'approche développée p. 39

I.4 Organisation de l'étude p. 48

Deuxième Chapitre : Utilisation de STICS pour la préconisation de la fertilisation azotée p. 49

II.1 Introduction p. 51

II.2 Utilisation d'un modèle de culture pour préconiser des doses optimales d'azote : revue bibliographique p. 52

II.3 Présentation du modèle de culture et des données utilisées p. 65

II.4 Paramétrage de STICS et évaluation de sa réponse à la fertilisation azotée p. 88

II.5 Choix d'un critère agro-environnemental p. 96

II.6 Evaluation de l'aptitude de STICS à déterminer un stratégie de fertilisation azotée selon un critère agro-environnemental p. 114

II.7 Conclusion p. 125

Troisième Chapitre : Elaboration de deux approches de spatialisation du modèle de culture à l'échelle intra-parcellaire p. 127

III.1 Introduction p. 129

III.2 Spatialisation du modèle basée sur un inventaire cartographique des sols p. 130

III.3 Spatialisation du modèle par assimilation de données p. 157

III.4 Conclusion p. 221

Quatrième Chapitre : Mise en œuvre de l'outil de préconisation spatiale de la fertilisation azotée p. 223

IV.1 Introduction p. 225

IV.2 Mesures spatialisées de variables caractéristiques de la culture p. 226

IV.3 Comparaison de l'approche par cartographie et de l'approche par assimilation p. 245

IV4 Mise en œuvre expérimentale de la méthode p. 272

IV.5 Conclusion p. 280

Cinquième Chapitre : Discussion et conclusion p. 281

I - Premier chapitre
Contexte de l'étude, objectifs et démarche

Si les créations de la science des hommes nous épatent,
c'est que nous sommes faciles à épater.

Jean Giono

I.1 Introduction 25
I.2 Fertilisation azotée et agriculture de précision 26
I.2.1 Bref aperçu du contexte social et politique 26
I.2.2 La nécessité de raisonner les itinéraires techniques à l'échelle intra-parcellaire 27
I.2.3 L'agriculture de précision, apparition et développement actuel 28
I.2.4 Deux grandes approches pour la gestion spatiale de la fertilisation azotée 29
I.2.5 Spatialisation des données : krigeage et télédétection 36
I.3 Présentation de l'approche développée 39
I.3.1 Choix du modèle de culture 39
I.3.2 Définitions et formalismes du modèle 40
I.3.3 Prise de décision sur la fertilisation azotée par le modèle 41
I.3.4 Spatialisation des données d'entrée du modèle de culture 42
I.3.5 Brève description des jeux de données utilisées 45
I.4 Organisation de l'étude 48

I.1 Introduction

Ce travail a été initié au sein d'un projet développé dans le cadre de l'Action Incitative Programmée « Agriculture de Précision » de l'INRA impliquant les équipes Agronomie de Laon, Science du Sol d'Orléans, Climat, Sol et Environnement et Biométrie d'Avignon. Le but de ce projet était de mettre en place des méthodes permettant d'adapter un certain nombre de types d'itinéraires techniques à l'hétérogénéité qui existe au sein d'une parcelle.

Cette étude concerne la fertilisation azotée du blé d'hiver à l'échelle intra-parcellaire. Elle consiste à mettre au point un outil, basé sur un modèle de culture et l'utilisation de données issues de la télédétection, permettant de prendre en compte (i) la variabilité des sols d'une parcelle agricole et (ii) les interactions du sol avec la culture et le climat afin d'adapter les apports d'azote minéral aux besoins locaux de la parcelle. Le raisonnement de la quantité de fertilisant apportée se fera de façon à optimiser plusieurs variables pouvant être antagonistes : le rendement, la qualité des grains via la teneur en protéines et l'impact sur l'environnement. Le support expérimental principal de notre étude correspond à deux parcelles utilisées dans le cadre du projet agriculture de précision de l'INRA, situées non loin de Laon, dans l'Aisne, en Picardie, dans une exploitation agricole (Guérif *et al.*, 2001b). La fréquence de retour du blé étant de deux ans, un suivi de ces parcelles a permis d'établir un jeu de données de quatre années-parcelles, de la saison culturale 1999-2000 à 2002-2003. Au cours des deux premières années, des apports d'azote homogènes sur les parcelles ont été effectués, ce qui a permis de caractériser la variabilité spatiale de la culture liée à l'hétérogénéité du sol ; les deux dernières années ont fait l'objet d'expérimentations sur la modulation des apports.

Ce premier chapitre introductif permet de définir le sujet dans son contexte social et scientifique : réduction des intrants dans le but de diminuer l'impact de l'agriculture intensive sur l'environnement et développement de l'agriculture de précision. Elle décrit ensuite deux approches possibles pour raisonner la fertilisation azotée à l'échelle intra-parcellaire : la première, basée sur l'utilisation d'indicateurs du statut azoté de la culture ou du sol, qui est actuellement utilisée de façon pré-commerciale en France et la seconde, basée sur l'emploi de modèles de culture, qui fait précisément l'objet de notre étude. Les différentes modalités de mise en œuvre de cette dernière seront alors détaillées plus amplement, ce qui permettra de décrire l'organisation de la présentation de ce travail.

I.2 Fertilisation azotée et agriculture de précision

Cette première partie décrit (i) le contexte dans lequel est apparue la notion d'agriculture de précision, (ii) ce en quoi elle est capable de répondre à la volonté sociale et politique de minimiser les effets négatifs de l'agriculture intensive sur l'environnement et (iii) ses limites actuelles. Puis deux grandes approches de la fertilisation azotée à l'échelle intra-parcellaire seront définies.

I.2.1 Bref aperçu du contexte social et politique

Face à une demande sociale croissante, des contraintes environnementales se sont mises en place en Europe occidentale afin de limiter l'impact de l'agriculture intensive sur la pollution des eaux ou de l'air par le biais de l'emploi massif de fertilisants et de pesticides notamment. L'exemple le plus caractéristique concerne la pollution des nappes phréatiques par les nitrates et la volonté de limiter leur concentration à 50 mg l^{-1} dans les eaux de consommation. Or, il a été montré que l'adaptation de la fertilisation azotée aux besoins de la culture permet de diminuer la quantité d'azote restant dans le sol à la récolte et qui est susceptible, notamment pour le blé, d'être lixiviée et emportée vers le sous-sol au cours de l'hiver suivant (*e.g.* Jeuffroy & Recous, 1999 ; Mary *et al.*, 2002).

Par ailleurs, l'augmentation de la population mondiale, la diminution de la surface agricole utile (due notamment à l'accroissement de la taille des villes) et la fluctuation des coûts des intrants prêchent pour un meilleur raisonnement des itinéraires techniques dans le but d'augmenter l'efficience de l'utilisation des sols et des intrants, d'un point de vue agronomique et économique (Morgenthaler *et al.*, 2003 ; Seelan *et al.*, 2003).

Or, l'agriculture de précision consiste à adapter, à l'échelle intra-parcellaire, les intrants et les pratiques culturales aux besoins locaux de la culture de façon à optimiser la production. On résume le plus souvent cet objectif par la formule : « pratiquer la bonne intervention au bon endroit et au bon moment » (Zwaenepoel & Le Bars, 1997). Ceci peut s'appliquer à tous les aspects de l'itinéraire technique : fertilisation, fumure organique, gestion des adventices, des maladies et des ravageurs, travail du sol, irrigation… A partir de ce constat, de nombreux auteurs voient dans cette forme d'agriculture une réponse à la double problématique, environnementale et économique, présentée ci-dessus (Engel, 1997 ; Paz *et al.*, 1999 ; Schröder *et al.*, 2000 ; Booltink *et al.*, 2001 ; Godwin *et al.*, 2003 ; Morgenthaler *et*

al., 2003 ; Welsh *et al.*, 2003). Zhang *et al.* (2002) la rapprochent explicitement du concept d'agriculture durable.

I.2.2 La nécessité de raisonner les itinéraires techniques à l'échelle intra-parcellaire

L'adaptation des intrants, en terme de quantité et de dates d'apport, dans le but de satisfaire la demande de la culture et de diminuer les pertes (Jeuffroy & Recous, 1999), est du ressort de ce que l'on appelle l'agriculture raisonnée. L'agriculture de précision peut être vue en première approche comme une application de ces principes à l'échelle intra-parcellaire.

L'émergence de nouvelles techniques à l'origine du concept de l'agriculture de précision a permis de mettre en évidence l'hétérogénéité des parcelles agricoles de façon spatialement continue (*e.g.* Welsh *et al.*, 2003). L'imagerie aérienne, technique ancienne permettant d'observer les cultures mais peu employée auparavant (Batchelor *et al.*, 2002) a été relayée par la télédétection satellitale ou aérienne (*e.g.* Moulin *et al.*, 1998 ; Basso *et al.*, 2001). La télédétection correspond à l'ensemble des techniques permettant d'observer des caractéristiques de la terre à l'aide de capteurs embarqués sur des avions ou des satellites qui mesurent l'énergie émise par les surfaces observées. L'avènement des cartes de rendement a cependant réellement marqué le début de la sensibilisation des agriculteurs et des scientifiques à la réalité de l'hétérogénéité du couvert dans les parcelles (Paz *et al.*, 1999). L'établissement de cartes de rendement a été rendu possible par l'apparition (i) des systèmes de positionnement géographique (GPS, *Global Positioning System*) qui permettent de connaître en temps réel les coordonnées géographiques d'une machine, et (ii) de capteurs de rendement qui mesurent en continu le flux de grains récoltés dans la moissonneuse-batteuse. Cette hétérogénéité à la récolte est liée à plusieurs facteurs : la topographie de la parcelle, la variabilité de la fertilité du sol, de l'implantation de la culture, des attaques d'adventices et parasitaires, ainsi que de l'application de l'itinéraire cultural choisi, même à consigne homogène (Zhang *et al.*, 2002). A ce titre, Makowski *et al.* (2000) ont montré que l'impact économique et environnemental d'un épandage hétérogène d'engrais azoté n'était pas négligeable. Jusqu'à une période récente, cette hétérogénéité n'était pas prise en compte pour plusieurs raisons. La plus faible taille des parcelles, parfois délimitées sur des critères d'homogénéité, faisait que cette variabilité intra-parcellaire était elle-même de plus faible ampleur. L'absence de moyens de caractérisation de cette variabilité et de sa prise en compte dans le raisonnement des itinéraires techniques conduisait à pratiquer des conduites de culture uniformes (Zwaenepoel & Le Bars, 1997). Aujourd'hui, l'augmentation des surfaces des

exploitations et le remembrement d'une part et l'avènement des capteurs de rendement d'autre part ont rendu nécessaire et possible cette prise en compte.

L'agriculture de précision permet de raisonner les pratiques culturales à l'échelle intra-parcellaire grâce à l'emploi de Systèmes d'Information Géographiques (SIG) qui permettent de gérer des bases de données spatiales. Dans les développements actuels les plus fréquents, cette gestion de l'hétérogénéité consiste à identifier au sein des parcelles des zones homogènes qui peuvent être gérées selon des méthodes classiques (Engel, 1997 ; Batchelor *et al.*, 2002 ; Zhang *et al.*, 2002 ; Seelan *et al.*, 2003 ; Welsh *et al.*, 2003).

I.2.3 L'agriculture de précision, apparition et développement actuel

L'agriculture de précision s'est développée en premier lieu aux Etats-Unis au début des années 1980, alors que des analyses de sol en plusieurs lieux de la parcelle permettaient une modulation des apports d'engrais. Mais c'est l'apparition des cartes de rendement au début des années 1990 qui a réellement marqué l'essor du concept de l'agriculture de précision (Zwaenepoel & Le Bars, 1997) et, dans une moindre mesure, d'appareils permettant de moduler les pratiques culturales, au premier rang desquels des épandeurs d'engrais à débit piloté. Ceci illustre le fait qu'historiquement, l'émergence de cette forme d'agriculture est essentiellement le fait du développement des technologies décrites ci-dessus. Pierce & Nowak (1999) détaillent les conditions qui ont permis son essor : le développement de l'informatique, l'apparition des SIG, des GPS, de capteurs permettant de caractériser de façon géographiquement continue l'état de la culture ou du sol et de machines permettant de mettre en œuvre des techniques culturales de façon spatialement modulée. Guérif *et al.* (2001) proposent de distinguer trois phases distinctes dans la mise en œuvre de l'agriculture de précision (Figure I.2-1), qui doivent être gérées de façon informatique (utilisation d'un SIG) : (i) la caractérisation de la variabilité spatiale de la parcelle ; (ii) la prise en compte de cette variabilité pour préconiser des cartes ; (iii) la mise en œuvre de la décision spatialement variable. Le premier et le dernier points correspondent essentiellement aux technologies qui sont, on l'a vu, à l'origine de l'agriculture de précision. Le deuxième aspect est resté le plus longtemps déficient, ce qui fait dire à Zhang *et al.* (2002), de concert avec nombre d'autres auteurs (Zwaenepoel & Le Bars, 1997 ; Pierce & Nowak, 1999 ; Robert, 2000 ; Seelan *et al.*, 2003), que, si les conditions d'innovations technologiques de la mise en œuvre de l'agriculture de précision sont à présent réunies, le développement de principes agronomiques et écologiques pour optimiser les recommandations spatiales des techniques culturales reste

encore très généralement partiel. On se trouve très souvent dans des situations où l'on ne sait pas comment utiliser l'information, ce qui est typiquement le cas des cartes de rendement : l'agriculture de précision doit donc être, plus que l'injection de nouvelles technologies dans le monde agricole, une nouvelle forme d'acquisition et de gestion de l'information (Robert, 2000). De ce point de vue, elle en est encore à ses balbutiements (Pierce & Nowak, 1999 ; Zhang *et al.*, 2002). Malgré l'évidence intuitive des apports potentiels de l'agriculture de précision en termes économiques et environnementaux (Robert, 2000 ; Godwin *et al.*, 2003 ; Morgenthaler *et al.*, 2003), rien n'a encore été démontré de manière irréfragable (Zhang *et al.*, 2002 ; Welsh *et al.*, 2003 ; James & Godwin, 2003) et ce d'autant plus que l'intérêt de la gestion spatiale des techniques dépend lui-même du degré de variabilité spatiale de la parcelle (Godwin *et al.*, 2003). Pour toutes ces raisons, sans compter l'investissement financier important et difficilement chiffrable nécessité par ces techniques, de nombreux agriculteurs hésitent à franchir le pas.

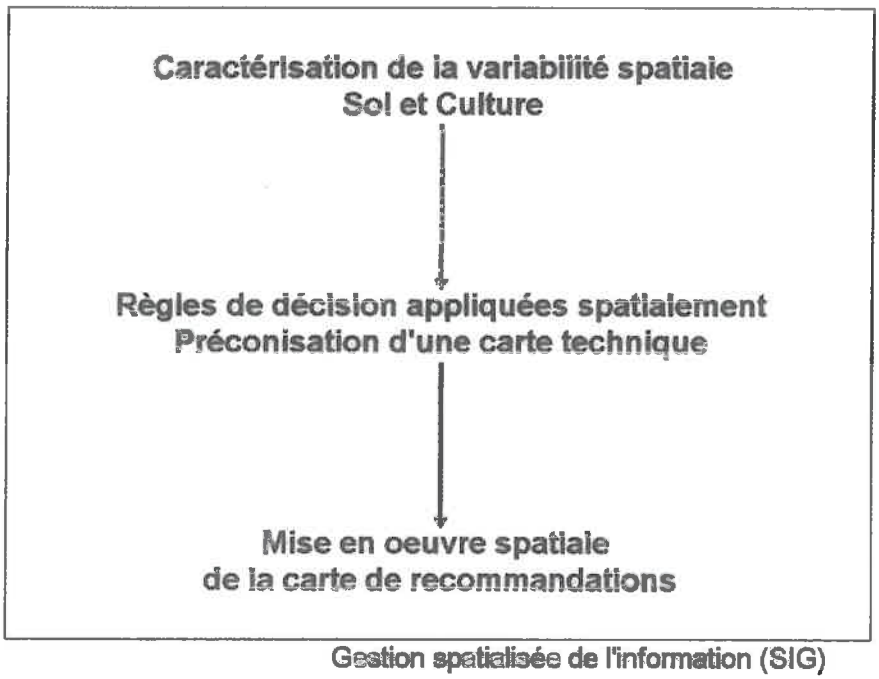


Figure I.2-1 : Les trois principales phases de la mise en œuvre des principes de l'agriculture de précision. D'après Guérif *et al.* (2001).

I.2.4 Deux grandes approches pour la gestion spatiale de la fertilisation azotée

La vision bibliographique rapide des tenants et aboutissants actuels de l'agriculture de précision qui vient d'être faite nous conduit à restreindre à présent le champ de notre étude.

Dans le cadre de notre étude, nous nous attacherons uniquement aux deux premières étapes décrites dans la Figure I.2-1 pour la gestion de la fertilisation azotée : la description de la variabilité spatiale du système sol-plante et la mise en place de règles de décision permettant d'aboutir à la préconisation des cartes d'épandage. Seules les doses à prescrire seront prédites, les dates d'apports ne seront pas non plus gérées par le système, bien que ce soit potentiellement possible. La mise en œuvre des recommandations spatiales sera prise en charge par le CEMAGREF. Le problème des attaques parasitaires ne sera pas abordé, ainsi que celui des carences en éléments chimiques autres que l'azote (phosphore et potassium notamment). Seuls les facteurs limitants liés à la fertilité du sol en azote et à l'eau seront donc pris en compte. L'intérêt économique de la mise en œuvre de l'agriculture de précision dans notre situation d'étude ne sera pas abordé, non plus que l'intérêt de réaliser des modulations, ce qui serait du reste possible dans un proche avenir.

En région de grande culture comme la Picardie, les apports d'azote se font de façon assez générale en trois ou quatre apports. Le premier, en sortie d'hiver (fin février, début mars) est assez peu important (environ 40 kg ha^{-1}) ; le dernier (début mai) sert à assurer une bonne teneur en protéines des grains. Les apports qu'il est le plus important de raisonner sont les plus conséquents en quantité, qui ont lieu vers la fin mars, début avril. Cela implique que les méthodes de raisonnement de la fertilisation azotée soient capables d'établir à cette période l'état azoté de la culture.

Le raisonnement de la fertilisation azotée dans le cadre de l'agriculture de précision peut se faire en spatialisant les méthodes de préconisation existantes. Actuellement, les méthodes opérationnelles sont basées sur des indicateurs caractéristiques du sol (*e.g.* AZOBIL, Machet *et al.*, 1990) ou de la culture (*e.g.* JUBIL, Justes *et al.*, 1997). Des méthodes basées sur l'utilisation de modèles de culture sont également développées (Engel, 1997 ; Shaffer & Brodahl, 1998 ; Jeuffroy & Recous, 1999 ; Booltink *et al.*, 2001 ; Loyce *et al.*, 2002 ; Jones *et al.*, 2003). Nous reprendrons la même dichotomie pour décrire deux grands types d'approches applicables en agriculture de précision :

- l'approche « indicateur », qui généralise à l'échelle intra-parcellaire des principes développés à l'échelle parcellaire ;

- l'approche « modèle de culture » qui est l'objet de notre étude.

Ces deux approches sont illustrées par la Figure I.2-2 pour la première et la Figure I.2-3 pour la seconde.

I.2.4.1 Description de l'approche indicateur

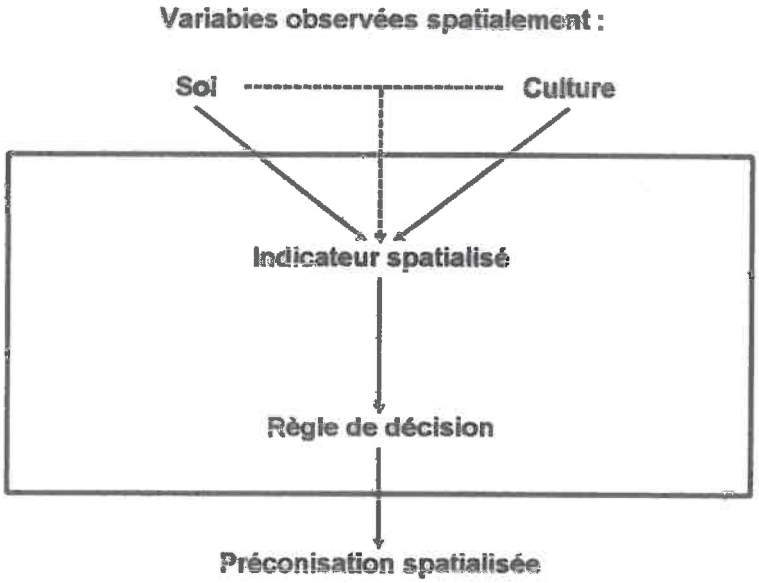


Figure I.2-2 : Principe de l'approche « indicateur » permettant de spatialiser à l'échelle intra-parcellaire des préconisations de fertilisation azotée. Il ne prend en compte que des observations limitées en nombre dans le temps de variables caractéristiques du sol et de la culture. Les deux types d'information peuvent être employés individuellement (flèches pleines) ou de façon combinée (flèche en pointillé) pour déterminer un indicateur qui servira à réaliser les préconisations spatiales via une règle de décision.

Pierce & Nowak (1999) distinguent trois méthodes possibles de la gestion intra-parcellaire de la fertilisation azotée :

- (i) utiliser la méthode du bilan (Stanford, 1973) de façon spatialisée ;
- (ii) déterminer des courbes de réponses à la dose d'azote et la dose optimale associée en fonction du type de sol ;
- (iii) déduire la dose d'azote du statut azoté de la culture estimé spatialement.

Ces trois façons de faire impliquent de mesurer de façon spatialisée, une ou plusieurs fois au cours du cycle, une grandeur – l'indicateur – qui permet de prendre la décision (Figure I.2-2). Notons tout de suite que la deuxième de ces méthodes est peu susceptible d'être appliquée en pratique car l'établissement de courbes de réponse à l'échelle de la parcelle n'est pas une mince affaire et, par ailleurs, elles sont susceptibles de varier d'une année sur l'autre. Sur le plan expérimental en revanche, elle est utile pour mettre en évidence l'hétérogénéité

intra-parcellaire du point de vue de la réponse de la culture à la dose d'azote (Bramley *et al.*, 1999 ; Laurent, 2000). On peut aussi proposer de classer les différentes méthodes s'apparentant à l'approche indicateur selon la nature de cette grandeur : un indicateur caractéristique du sol, ou de la culture. Nous allons donner un bref aperçu des approches de type indicateur selon cette distinction.

Historiquement, les indicateurs caractéristiques du sol ont été les premiers utilisés : des prélèvements de sol effectués selon une grille systématique ou orientée par des observations permettent par analyse chimique de mettre en évidence des carences de certains éléments pour tous les points de mesure. Les règles de préconisation classiquement utilisées peuvent ensuite être appliquées. On peut par exemple citer une expérimentation (Machet, non publié) qui a consisté à spatialiser les préconisations de la méthode AZOBIL (Machet *et al.*, 1990) à l'échelle intra-parcellaire. Cette méthode est basée sur le concept de balance azotée (Stanford, 1973) qui détermine la « dose bilan » à apporter en calculant la différence entre les besoins de la culture et les apports susceptibles d'être réalisés par le sol, par la minéralisation de la matière organique ou par le biais des reliquats d'azote minéral. Les besoins de la culture sont déterminés par un rendement objectif qui peut être lui même spatialisé en fonction des rendements obtenus les années précédentes ou d'une estimation en cours de culture : nous rejoignons ici les indicateurs caractéristiques de la culture.

L'utilisation en agriculture de précision d'indicateurs caractéristiques de la culture permet aussi de généraliser des techniques préexistantes. Huet *et al.* (2004, à paraître) proposent de reprendre un principe de fertilisation du colza basé sur l'estimation de la biomasse en sortie d'hiver. Dans le cas de l'utilisation à l'échelle intra-parcellaire, la biomasse est estimée de façon spatialement continue par l'intermédiaire de mesures de télédétection. Booij & Uenk (2001) utilisent également des mesures de télédétection pour estimer empiriquement la quantité d'azote absorbé par la culture et en déduisent le déficit d'absorption. Akkal (1998) utilise des mesures de télédétection en sortie d'hiver pour déterminer, *via* l'estimation de l'indice de surface foliaire (*LAI*), un potentiel de production de biomasse, et en déduire les besoins en azote. Le lien entre biomasse et besoin en azote se fait grâce au concept de courbe du dilution (Lemaire & Gastal, 1997). Enfin, Gate (2000) propose de combiner les deux approches : en cours de culture, des mesures de télédétection permettent d'estimer à la fois le statut azoté de la culture et son développement potentiel : en utilisant le concept de l'indice de nutrition azoté, ou INN, (Lemaire & Gastal, 1997), on peut en déduire le déficit d'absorption au moment de l'observation.

Les deux types d'indicateurs (sol et plante) peuvent également être combinés. A l'échelle de la parcelle entière, la méthode JUBIL (Justes *et al.*, 1997) consiste premièrement à déterminer la dose bilan grâce à la balance azotée. Cette dose n'est pas entièrement appliquée. Des mesures du statut azoté (teneur en nitrates du jus de base de tige, corrélé avec l'indice de nutrition azotée) permettent à des stades de croissance donnés de déterminer si le blé est carencé ou pas. Si c'est le cas, la part de la dose laissée en réserve est appliquée. La méthode FARMSTAR de la société Astrium, utilisée de façon pré-commerciale, consiste à remplacer la mesure JUBIL par une estimation du statut azoté *via* la télédétection. En revanche, seule la dose complémentaire est spatialisée par cette méthode, la détermination de la dose bilan ne l'est pas (Gate, comm. pers.).

Pour appliquer ces méthodes pré-existantes à l'échelle intra-parcellaire, on peut procéder soit à un krigeage des mesures ou des préconisations effectuées (McBratney & Pringle, 1999), soit à un zonage de la parcelle grâce à des cartes de sol ou à des observations de télédétection par exemple. Dans le second cas, la méthode est appliquée pour chaque petite zone. Nous définirons ultérieurement les termes de télédétection et de krigeage comme des techniques de spatialisation.

L'intérêt de l'approche par indicateur est qu'elle utilise des méthodes robustes et éprouvées. Certaines d'entre elles, à l'instar de JUBIL ou des ses homologues spatialisées, combinent une caractérisation du sol et de la culture, ce qui constitue un atout indéniable par rapport à celles qui n'en utilisent qu'une : en effet, un état de carence de la plante peut par exemple être lié à une faible minéralisation temporaire. Si les conditions climatiques changent et favorisent à nouveau la minéralisation, il serait préjudiciable d'avoir apporté de l'azote minéral.

Quoiqu'il en soit, aucune de ces méthodes ne permet de prendre en compte explicitement l'évolution temporelle du système sol-culture de façon continue : les observations sont très ponctuelles, la plupart du temps à des stades clefs du développement de la culture. De ce fait, elle ne prennent pas non plus en compte le climat de l'année, qui influence notamment le développement de la culture, et donc ses besoins, et la minéralisation de l'azote dans le sol. Certaines d'entre elles utilisent cependant, pour passer des estimations d'indice de surface foliaire à la biomasse produite, des modèles de croissance (Akkal, 1998 ; Gate, 2000). Ceci illustre l'impérieuse nécessité de prendre en compte le climat pour adapter au mieux les apports aux besoins. Mais les modèles utilisés dans ce cadre ne font que traduire

une croissance potentielle, c'est-à-dire en absence de stress. Les carences en azote elles-mêmes ne sont donc pas simulées. Enfin, ces méthodes ont été essentiellement développées dans le but d'optimiser une seule variable, le rendement, et ne permettent donc pas de prendre en compte des facteurs environnementaux et de qualité de la récolte.

I.2.4.2 Description de l'approche modèle de culture

La Figure I.2-3 représente l'approche basée sur l'utilisation d'un modèle de culture développée dans ce travail.

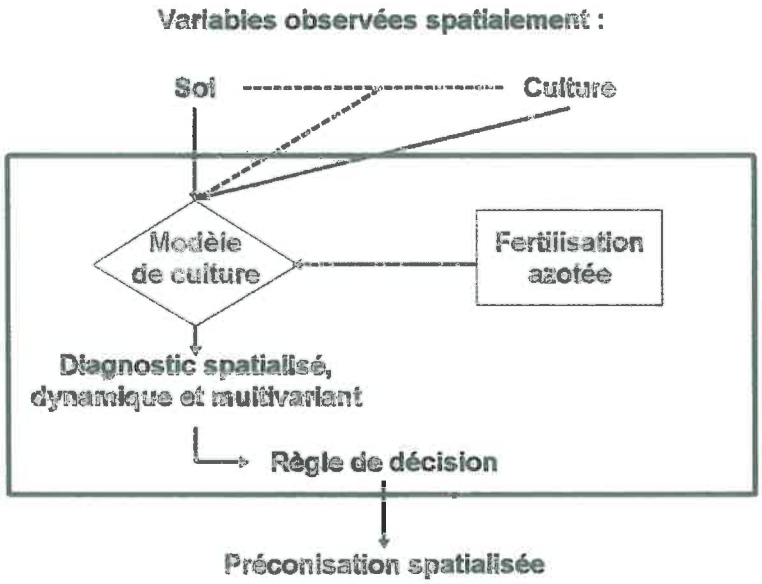


Figure I.2-3 : Principe de l'approche modèle de culture : deux types de variables d'entrée sont susceptibles de servir à représenter spatialement à l'aide du modèle l'état de croissance de la culture, individuellement (flèches pleines) ou conjointement (flèche en pointillé). Grâce à sa description de la réponse du couvert à la fertilisation azotée, le modèle peut (i) porter un diagnostic spatialisé, dynamique et multivariant du statut azoté de la culture et (ii) servir à réaliser une préconisation spatiale de la fertilisation.

Un modèle de culture représente sous forme d'équations le développement et la croissance d'une culture. Ces équations synthétisent des connaissances sur des processus écophysiologiques, et sur les interactions qui existent entre la culture, le climat et le sol.

Par rapport à l'approche qui vient d'être présentée, celle qui est basée sur de tels modèles comporte plusieurs intérêts (Guérif *et al.*, 2001a) :

- les caractéristiques du sol, de la plante et du climat de l'année sont formellement prises en considération : les informations d'entrée sont multi-variables. Si ces entrées sont renseignées de façon spatiale, le modèle peut simuler la croissance de la culture en tout point de la parcelle ;

- l'aspect dynamique est formalisé et les interactions entre différents stress sont simulées, ce qui permet d'établir un diagnostic spatialisé, dynamique et multivariant du statut azoté de la culture ;

- plusieurs variables de sortie du système sol-plante sont simulées qui permettent d'élaborer des critères agro-environnementaux prenant en compte à la fois des aspects de rendement, de qualité et de risques pour l'environnement ;

- le modèle a des vertus prédictives, ce qui permet de prendre en compte des risques climatiques et de tester de nombreux scénarios d'itinéraires techniques.

Pour pouvoir l'utiliser dans le cadre de l'agriculture de précision, il faut s'assurer que le modèle puisse simuler la croissance de la culture à l'échelle intra-parcellaire. Pour ce faire, deux méthodes de prise en compte de la variabilité intra-parcellaire, appelées méthodes de spatialisation, sont envisageables, de façon complémentaire ou non :

- la première consiste à renseigner les entrées nécessaires au modèle en caractérisant par des mesures les propriétés du sol à l'échelle intra-parcellaire ;

- la seconde utilise des mesures spatialisées des caractéristiques du couvert, par le biais de la télédétection par exemple, pour estimer les données d'entrée spatiales du modèle.

Le modèle doit également permettre de décrire la réponse du système sol-plante à différentes doses d'azote, afin de pouvoir l'utiliser pour réaliser des préconisations selon des règles de décision.

L'utilisation d'un modèle de culture pour réaliser des préconisations spatiales de fertilisation azotée est abordée par de nombreux auteurs : la plupart d'entre eux utilisent des mesures du sol pour renseigner les variables d'entrée de ces modèles, ce qui est la voie la plus naturelle (Engel, 1997 ; Paz *et al.*, 1999 ; Batchelor *et al.*, 2002 ; van Alphen, 2002 ; Nicoullaud *et al.*, 2003). Très souvent, comme dans le cas de l'approche indicateur, ces mesures sont effectuées de façon ponctuelle et ensuite krigées.

Certains auteurs utilisent des images de télédétection pour délimiter des zones « homogènes » pour réaliser les simulations (Basso *et al.*, 2001 ; Batchelor *et al.*, 2002). Des informations venant du sol lui-même ou des cartes de rendement peuvent également servir à

définir ces zones homogènes de gestion de la fertilisation (Engel, 1997). La définition de telles zones permet notamment de diminuer le nombre de mesures à réaliser pour estimer les variables d'entrée des modèles dans la parcelle car, comme le souligne Engel (1997), leur coût est prohibitif. Certaines d'entre elles sont même très difficiles à estimer, comme la profondeur d'enracinement potentiel (Batchelor *et al.*, 2002), ce qui peut conduire à utiliser les données de rendement estimées par des cartes pour ajuster les valeurs de ces variables en minimisant la différence entre les simulations et les observations sur les zones de sol définies. La tendance est donc, comme le suggère Morgenthaler *et al.* (2003), d'utiliser le maximum d'informations possibles.

Dans le cadre de notre travail, nous avons utilisé l'approche modèle de culture en cherchant à utiliser à la fois des informations provenant de mesures des caractéristiques du sol et des variables biophysiques issues de la télédétection selon des principes avancés par certaines équipes et que nous décrirons plus avant dans la partie suivante.

I.2.5 Spatialisation des données : krigeage et télédétection

Les deux types d'approches définies ci-dessus utilisent deux voies de renseignement de leurs données d'entrée lorsqu'elles sont employées de façon spatiale :

- des mesures effectuées de façon ponctuelle et ensuite généralisées à l'ensemble de la parcelle par krigeage ;

- des mesures de télédétection qui permettent d'emblée d'avoir une caractérisation spatialement continue de l'espace.

Ces deux termes sont définis dans cette partie.

I.2.5.1 Définition simple du krigeage

Il s'agit d'une technique de géostatistique qui permet d'interpoler des mesures faites de façon très localisée à l'ensemble de l'espace. La corrélation des mesures dans un voisinage géographique est utilisée pour définir un estimateur local de la grandeur mesurée là où la mesure n'a pas été faite. Cet estimateur est une régression multiple des données mesurées (Chilès & Delfiner, 1999). C'est une technique couramment utilisée, comme on l'a vu, pour établir des cartes de caractéristiques de sol notamment.

I.2.5.2 Notions élémentaires de télédétection et applications courantes en agriculture de précision

La télédétection consiste à mesurer la réflectance de surfaces observées grâce à des capteurs embarqués sur des satellites ou des avions. La réflectance est le rapport entre la quantité de lumière réfléchie par la surface et la lumière incidente. Pour la caractérisation d'un couvert végétal, les mesures sont effectuées essentiellement dans le domaine du visible (400-780 nm) et du proche infra-rouge (780-900 nm). A partir de ces mesures de réflectance, deux grandes voies sont utilisées pour calculer des variables biophysiques caractéristiques du couvert comme l'indice de surface foliaire (*LAI*) et la teneur en chlorophylle (*Cab*) :

- le calcul d'indices de végétation, combinaison de réflectances à des longueurs d'onde données. Ces valeurs d'indice sont corrélées aux caractéristiques du couvert, par exemple le « red-edge » (point d'inflexion de la courbe de la réflectance en fonction de la longueur d'onde situé dans le domaine du rouge) est corrélé à *Cab* et le NDVI (combinaison des réflectances dans le rouge et le proche-infrarouge) au *LAI* ;

- l'utilisation de modèles de transfert radiatif, qui simulent selon des lois physiques la relation entre (i) les propriétés du couvert et les caractéristiques des configurations d'éclairement et d'observation et (ii) les réflectances à différentes longueurs d'onde. Par inversion de ces modèles, c'est-à-dire l'estimation des entrées en connaissant les sorties, on peut connaître notamment *LAI* et *Cab* qui font partie des entrées.

Par rapport à notre problématique de la fertilisation azotée, nous exploiterons le fait qu'il existe un lien fort entre teneur en chlorophylle et teneur en azote du couvert (*e.g.* Dumont & De Baerdemaeker, 2001). Nous y reviendrons plus largement en temps utile.

Dans les travaux utilisant la télédétection en agriculture de précision, les indices de végétation sont encore majoritairement utilisés (*e.g.* Vouillot *et al.*, 1998 ; Booij & Uenk, 2001 ; Dumont & De Baerdemaeker, 2001 ; Batchelor *et al.*, 2002 ; Welsh *et al.*, 2003). Nous utiliserons dans cette étude les variables biophysiques *LAI* et *Cab* issues de l'inversion des modèles de transfert radiatif telle que décrite par Baret & Fourty (1997) et Moulin & Guérif (soumis).

Les avantages de la télédétection dans le cadre de l'agriculture de précision sont les suivants (Guérif *et al.*, 2001a) :

- l'accès à des variables d'état du couvert en lien avec leur fonctionnement, comme *LAI* et *Cab*, selon les capteurs utilisés ;

- l'exhaustivité spatiale de la mesure et la résolution compatible avec les impératifs de l'agriculture de précision grâce à différents vecteurs (satellites, avions, engins agricoles) ;

- la répétitivité temporelle de ces mesures et la potentialité de les utiliser en temps réel, qui devrait augmenter dans le futur.

Seelan *et al.* (2003) donnent quelques réflexions sur le fait que la télédétection est encore malgré tout peu employée en agriculture de précision alors qu'elle l'est beaucoup plus dans des applications à plus grande échelle. En substance, jusqu'à une période récente, les capteurs n'avaient pas la résolution, c'est-à-dire la taille du pixel élémentaire, et la fréquence d'observation requises ; l'agriculture de précision, contrairement à des gestions des cultures à grande échelle, s'adresse plus directement aux agriculteurs eux-mêmes qui n'ont pas forcément les compétences nécessaires pour extraire la bonne information de ces mesures. On retrouve ici le problème de gestion complexe de l'information déjà souligné à propos de l'utilisation des cartes de rendement.

Dans l'approche que nous avons utilisée, les données issues de la télédétection serviront à renseigner les variables d'entrée du modèle par des méthodes d'assimilation que nous allons définir dans la partie suivante.

I.3 Présentation de l'approche développée

Cette partie décrit dans un premier temps les caractéristiques requises par le modèle de culture afin de justifier le choix qui a été fait d'utiliser STICS pour cette étude. Elle donnera ensuite un certain nombre de définitions et de formalismes. Les deux grands volets impliqués par l'approche « modèle de culture » – (i) le choix de la dose et (ii) la spatialisation des simulations – seront ensuite développés. Enfin, un aperçu sera fait des jeux de données impliqués dans la mise en œuvre de ces volets.

I.3.1 Choix du modèle de culture

Dans le cadre de l'agriculture de précision, c'est surtout la variabilité des caractéristiques du sol qui est à l'origine de la variabilité de l'état du couvert. D'après la Figure I.2-3, le modèle doit avoir comme données d'entrée des informations concernant le sol. Ces informations doivent permettre de simuler, dans le modèle, les stress les plus fréquents qui limitent le développement potentiel de la culture : les stress azoté et hydrique.

La fertilisation azotée doit également être prise en compte par le modèle : plus précisément, la réponse de la culture à la fertilisation doit être simulée ainsi que les interactions entre les apports d'azote à la culture par le sol et par la fertilisation.

Du point du vue du climat, l'intérêt de l'utilisation d'un modèle par rapport à celle d'indicateurs est de pouvoir prendre en compte les spécificités du climat caractérisant la campagne de culture jusqu'à la date de prise de décision et de gérer l'incertitude sur le climat à venir en terme de risques. Le modèle choisi doit donc permettre de simuler l'influence des conditions climatiques sur le développement du blé. Cette influence doit être combinée à celle des propriétés du sol.

L'objectif de l'outil à développer est de limiter l'impact de la fertilisation azotée sur l'environnement tout en permettant d'atteindre de bons résultats en terme de rendement et de qualité des grains. Le modèle choisi doit donc simuler correctement ces variables.

Le modèle de culture sélectionné pour cette étude est le modèle STICS (Brisson *et al.*, 1998) qui a été développé pour simuler la croissance de plusieurs types de culture dont le blé en réponse à une grande variabilité de types de sols et de climats. Il inclut également l'impact de différentes pratiques culturales dont la fertilisation azotée et ses variables de sortie

comportent les principales variables d'intérêt agronomique, dont le rendement, la teneur en azote des grains et les reliquats d'azote dans le sol à la récolte. En outre, une analyse de sensibilité était disponible (Ruget *et al.*, 2002) et des travaux portant sur l'assimilation de données issues de la télédétection dans un modèle de culture l'avaient déjà utilisé (Weiss *et al.*, 2001 ; Prévot *et al.*, 2003).

I.3.2 Définitions et formalismes du modèle

Pour décrire un modèle de culture, nous adopterons le formalisme suivant : un modèle de culture est une fonction qui, à un ensemble de variables explicatives noté X et un ensemble de grandeurs que nous appellerons en première approche paramètres et noté θ, associe un vecteur de variables d'état simulées dépendant du temps t, noté $\hat{Y}(t)$:

$$\hat{Y}(t) = f(X;\theta;t) \qquad \text{Équation I.3-1}$$

On peut distinguer les variables d'état, qui sont simulées au pas de temps du modèle pour toute la durée de la simulation et les variables de sortie qui ne sont données qu'à la fin de la simulation. On peut écrire ces dernières : $\hat{Y}^S = f(X;\theta)$. Ce sont la plupart du temps des variables d'état considérées pour une valeur particulière de t, la récolte. Par rapport au problème qui nous préoccupe, nous considèrerons comme variables d'état essentiellement l'indice de surface foliaire (*LAI*) et la quantité d'azote absorbée par la culture (*QN*). Les variables de sortie que nous utiliserons le plus seront le rendement (noté G, pour rendement grain), la teneur en protéines des grains (P) et le reliquat d'azote dans le sol à la récolte (R). Comme ce sont ces variables qui serviront à calculer le critère permettant d'effectuer le choix de la fertilisation azotée et à le valider, nous les appellerons volontiers « variables d'intérêt final ».

Par paramètre, on entend généralement des valeurs qui ne sont pas modifiées au cours de la simulation, et par variable des valeurs qui sont susceptibles d'être recalculées à chaque pas de temps du modèle. Cependant, conformément à notre double problématique de gestion de la fertilisation azotée et de variabilité spatiale, nous utiliserons dans cette étude les acceptions suivantes :

- X correspond aux variables d'entrée du modèle exprimant le forçage climatique qui doivent être renseignées au pas de temps journalier du modèle. La notation $X^?$ indique que le climat est inconnu à partir de la date de choix de la fertilisation azotée ;

- θ correspond aux « paramètres » du modèle. Dans cet ensemble, nous distinguerons quatre sous-ensembles : (i) les paramètres qui ne seront jamais modifiés et que nous ignorerons pour cette raison ; (ii) les paramètres qui seront l'objet d'un paramétrage dans le Chapitre II : il s'agit essentiellement de paramètres concernant les caractéristiques de la culture et sont considérés constants dans le temps et l'espace. Ils seront désignés par « paramètres » ; (iii) les paramètres du modèle qui décrivent le sol : comme ces paramètres varient d'un endroit à l'autre de la parcelle et doivent donc être déterminés sur chaque entité de simulation, nous les appellerons désormais : « variables d'entrée » du modèle. Ce terme se retrouvera dans les Chapitres III et IV ; (iv) la fertilisation azotée, qui, dans l'optique de la modulation intra-parcellaire, sera également une « variable d'entrée » au sens géographique du terme : lorsqu'il s'agira de préconiser des doses d'azote, nous la mettrons en évidence en la notant N.

I.3.3 Prise de décision sur la fertilisation azotée par le modèle

La façon de déterminer par le modèle une dose d'azote optimale doit prendre en compte les contraintes de notre étude. Nous nous limitons à l'apport d'azote intermédiaire, le plus important en quantité, dont la décision doit être prise vers la fin du mois de mars ou le début du mois d'avril, ce qui implique que le climat après cette date est inconnu.

Le principe de choix de la dose d'azote consiste donc à déterminer la dose d'azote N qui optimise les variables d'intérêt final, $\hat{Y}^S$, en fonction de climats possibles $X^?$. L'intérêt d'un modèle de culture est précisément de pouvoir prédire la valeur de $\hat{Y}^S$ pour plusieurs combinaisons de N et de $X^?$. Pour choisir N, nous réaliserons donc des simulations croisant plusieurs hypothèses sur N et $X^?$.

La dose optimale, notée N^* dans le cas réel et $\hat{N}^*$ dans le cas où elle est déterminée par le modèle, sera choisie de façon à optimiser une fonction des trois variables d'intérêt final, $\hat{Y}^S$, le rendement, la teneur en protéines et une variable décrivant l'impact de la fertilisation sur l'environnement, notée E. Cette fonction sera nommée par la suite critère agro-environnemental.

Ces deux phases, (i) simulation de $\hat{Y}^S$ pour plusieurs hypothèses de N et de $X^?$, et (ii) détermination de $\hat{N}^*$ à partir de $\hat{Y}^S$, nécessitent en préalable d'évaluer la capacité du modèle à simuler correctement les variables d'intérêt final dans des situations où X et N sont connues. Le modèle sera donc évalué, après une phase de paramétrage, sur sa capacité à reproduire la

réponse du système sol-plante à une fertilisation azotée variable, et ce pour des conditions pédoclimatiques différentes. Le jeu de données utilisé pour ce faire devra être différent du jeu de données principal. Le critère agro-environnemental sera également déterminé grâce à ce jeu de données réelles.

Enfin, pour évaluer l'aptitude de STICS à faire des préconisations, les valeurs de $\hat{N}^*$ préconisées par STICS seront comparées aux valeurs de $\hat{N}^*$ données par une méthode de référence, AZOBIL (Machet *et al.*, 1990) ainsi qu'aux doses optimales vraies, N^*.

I.3.4 Spatialisation des données d'entrée du modèle de culture

Les deux façons de spatialiser les données d'entrée du modèle de culture que nous avons distinguées dans la première partie – (i) établir une carte des propriétés des sols et (ii) utiliser les données issues de la télédétection par assimilation – seront mises en œuvre sur les parcelles expérimentales et comparées en terme de qualité de la simulation de la variabilité spatiale de la culture. Ces deux sources sont d'ailleurs complémentaires : nous verrons que la cartographie des sols utilise des images de télédétection et que l'assimilation nécessite une information concernant le sol, qui peut être très générale et imprécise, nommée information *a priori* (IAP). Elle peut par exemple correspondre à la connaissance de la parcelle qu'a l'agriculteur, ou à des cartes des sols établies à petite ou moyenne échelle.

La façon d'établir des cartes des sols compatibles avec les impératifs de l'agriculture de précision sera décrite ultérieurement (Chapitre III). En revanche, nous donnons ici des notions essentielles de l'assimilation de données issues de la télédétection dans un modèle de culture.

L'assimilation de données est une pratique très courante dans les domaines de la météorologie, de l'océanographie et de l'hydrologie notamment. Elle reste encore très peu pratiquée dans les modèles de culture (Maas, 1988 ; Delécolle & Guérif, 1988 ; Bouman, 1992 ; Clevers & Leeuwen, 1996 ; Moulin *et al.*, 1998 ; Guérif & Dure, 2000 ; Weiss *et al.*, 2001 ; Launay, 2002 ; Prévot *et al.*, 2003). Très schématiquement, on peut distinguer trois grandes familles de méthodes d'assimilation, bien que d'autres classifications puissent être également adoptées (Delécolle *et al.*, 1992) :

- le forçage consiste à considérer les variables assimilées comme des entrées caractéristiques de l'année de simulation qui doivent donc être renseignées au pas de temps du modèle (et correspondent à X dans la notation adoptée) ;

- la correction des variables d'état observées : lors d'une observation, la variable d'état observée, par exemple le *LAI*, est corrigée en fonction de sa valeur observée, de sa valeur simulée et de la confiance relative qu'on attribue aux deux types de valeurs. Les autres variables d'état peuvent également être corrigées à l'aide des corrélations calculées entre ces variables. Entre deux observations, le modèle ainsi réinitialisé fonctionne normalement. L'exemple le plus connu de ce genre de méthodes est le filtre de Kalman ;

- l'estimation des variables d'entrée : ce type de méthodes, illustré par la Figure I.3-1, se rapproche beaucoup de l'inversion pratiquée par exemple sur les modèles de transfert radiatif. Elle consiste à déterminer les meilleures valeurs possibles de certaines des variables d'entrée qui permettent de minimiser la différence entre valeurs observées et simulées des variables servant à l'assimilation.

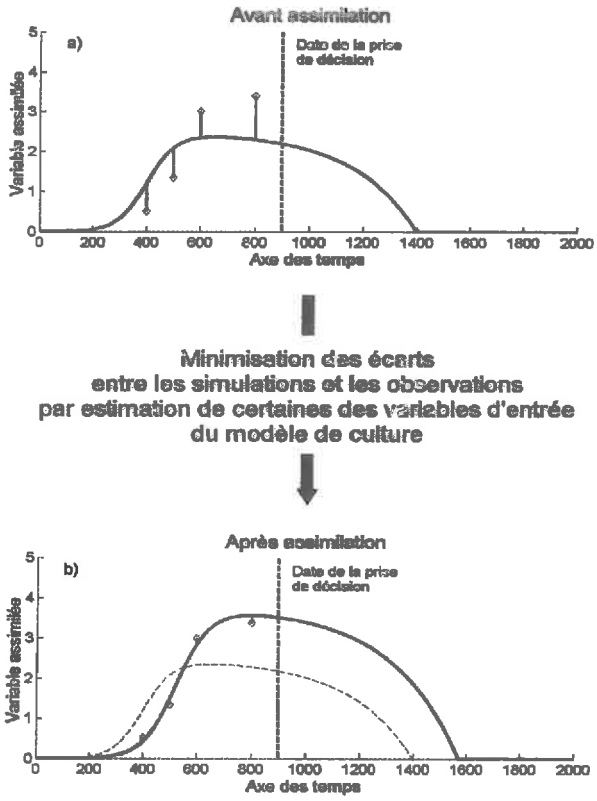


Figure I.3-1 : Assimilation de données par estimation des valeurs de certaines des variables d'entrée du modèle. Un algorithme calcule un ensemble de valeurs des variables d'entrée qui permet de minimiser la différence entre les valeurs observées et simulées par le modèle. Les dates d'observation doivent être positionnées avant la date de décision de la fertilisation azotée.

Nous avons choisi d'employer des méthodes estimant des variables d'entrée car elles répondent à l'objectif de caractériser la variabilité du sol, comme l'ont déjà fait Guérif & Duke (2000) et Launay (2002). Dans ce type de méthodes, les algorithmes les plus prometteurs sont ceux qui sont basés sur le calcul de la dérivée du modèle (méthode du gradient ou de l'adjoint), mais leur mise en œuvre est complexe et fait actuellement l'objet, pour STICS, d'une thèse (Lauvernet *et al.*, 2003).

Dans la plupart des applications aux modèles de culture que nous avons citées, le nombre de variables estimées est de fait limité par le nombre d'observations, ce qui implique de réaliser une analyse de sensibilité afin de déterminer quelles sont les variables qu'il est le plus pertinent de chercher à estimer. L'utilisation de ce que nous avons appelé l'information *a priori* permet de mettre en jeu davantage de variables.

L'information *a priori* (IAP) a deux fonctions : la première est de renseigner au mieux les variables d'entrée qui ne sont pas estimées par assimilation, de leur donner une valeur par défaut, ce qui est incontournable (*e.g.* Guérif & Duke, 2000 ; Launay, 2002 ; Prévot *et al.*, 2003). La seconde est d'affecter une distribution de valeurs probables aux variables que l'on estime. Les méthodes d'assimilation que nous utiliserons permettent de passer d'une distribution des valeurs des variables d'entrée, donnée par l'IAP, dite distribution *a priori*, à une distribution, dite *a posteriori*, qui intègre les informations apportées par les observations.

Le théorème de Bayes permet notamment ce passage de la distribution *a priori*, $P(\theta)$, à la distribution *a posteriori*, $P(\theta \mid Y)$, en utilisant les variables observées, Y (Makowski *et al.*, 2002) :

$$P(\theta|Y)=\frac{P(Y|\theta)P(\theta)}{P(Y)} \qquad \text{Équation I.3-2}$$

où $P(Y \mid \theta)$ représente la probabilité d'observer Y étant données les valeurs des variables d'entrée θ et $P(Y)$ est une constante de proportionnalité (déterminée par la nécessité que l'intégrale de $P(\theta \mid Y)$ sur l'ensemble de l'espace des paramètres soit égale à 1). Il sera utilisé par certaines des méthodes d'assimilation que nous mettrons en œuvre.

Trois méthodes d'assimilation seront comparées sur des jeux de données simulées dans le Chapitre III. Une d'entre elles sera utilisée sur le jeu de données de Chambry dans le Chapitre IV et comparée aux performances réalisées par la caractérisation par approche cartographique des entrées du modèle STICS.

L'utilisation de données de télédétection avec un modèle de culture pour faire de l'assimilation pose le problème de la correspondance entre les variables observées par télédétection et simulées par le modèle. En effet, les modèles de culture ne sont en général pas prévus pour simuler les réflectances observées par télédétection. Deux solutions sont envisageables :

- coupler le modèle de culture à un modèle de transfert radiatif, ce qui permet de créer un « sur-modèle » qui simule les sorties du modèle de transfert radiatif, les réflectances, à partir des entrées du modèle de culture ;

- utiliser les modèles de transfert radiatif pour les inverser de façon à estimer les variables biophysiques utiles au modèle de culture (*LAI* et *Cab* notamment) à partir des mesures de réflectances.

La plupart des travaux du type de celui auquel nous nous sommes attachés utilisent la première option (Guérif & Duke, 2000 ; Weiss *et al.*, 2001 ; Launay, 2002 ; Prévot *et al.*, 2003). Nonobstant l'intérêt de cette démarche, nous avons choisi la deuxième possibilité pour des raisons de simplicité d'une part, et parce que la précison de l'estimation de *LAI* et *Cab* à partir des réflectances s'est accrue (Moulin & Guérif, soumis). En outre, dans une perspective de comparaison de la méthode développée ici avec des approches de type indicateur, il était important de se situer sur le même plan car ces dernières utilisent elles-mêmes le produit des inversions des modèles de transfert radiatif.

Les deux solutions nécessitent de lier la teneur en chlorophylle, accessible par télédétection et la teneur en azote du couvert, simulée par les modèles de culture et utile pour le raisonnement de la fertilisation azotée. On peut par exemple lier *Cab* à la teneur en azote du couvert ou, comme nous le ferons, la quantité de chlorophylle présente dans le couvert (notée *QCab* et égale au produit de *LAI* et *Cab*) à la quantité d'azote absorbé par le couvert, noté *QN*.

I.3.5 Brève description des jeux de données utilisées

Les jeux de données seront plus largement présentés en temps voulu : le but de cette partie est de les nommer et de les situer dans l'étude.

I.3.5.1 Le jeu de données de Chambry

Dans le cadre de l'AIP Agriculture de Précision, deux parcelles expérimentales de 10 ha environ chacune ont été utilisées. Elles sont situées dans une exploitation agricole à Chambry, près de Laon, dans l'Aisne. Le Tableau I.3-1 présente pour chacune d'elles les saisons de culture au cours desquelles du blé de variété *Shango* y a été cultivé. Ces informations sont reprises sous forme de figure dans le marque page.

Saison	Parcelle	Dispositif expérimental	Engrais
1999-2000	1	Traitements azotés et acquisitions de télédétection	Liquide
2000-2001	2	Traitements azotés et acquisitions de télédétection	Liquide
2001-2002	1	Fertilisation azotée selon un carré latin	Solide
2002-2003	2	Fertilisation azotée modulée à l'est, constante à l'ouest	Solide

Tableau I.3-1 : Succession des cultures de blé dans les deux parcelles expérimentales de Chambry (02) et principaux aspects du dispositif expérimental.

En 1999-2000 et 2000-2001, six traitements de doses d'azote ont été réalisés et ont eu deux fonctions : (i) déterminer les relations entre *QCab* et *QN* de façon empirique et (ii) caractériser la réponse de STICS à la fertilisation azotée en combinaison avec les données présentées dans le paragraphe suivant. Des acquisitions de télédétection ont eu lieu au cours des quatre saisons, mais seules celles de 1999-2000 et de 2000-2001 seront utilisées dans ce travail. En 2001-2002, afin d'éprouver l'appareil servant à réaliser les modulations d'engrais et d'évaluer la capacité de STICS à représenter les interactions sol-fertilisation à l'échelle de la parcelle, la fertilisation a été apportée en trois doses disposées en carré latin. En 2002-2003, la méthode de préconisation spatialement variable a été testée : le côté Est de la parcelle a reçu un apport modulé, le côté Ouest un apport constant. La moyenne de cet apport était de 40 kg ha^{-1}. Au cours des deux premières campagnes, de l'engrais liquide a été employé car il permettait une meilleure homogénéité des apports. Mais il est caractérisé par une efficience plus faible. Lors des deux dernières campagnes, de l'engrais solide a été épandu car l'appareil utilisé pour réaliser les épandages modulés était adapté à ce type d'engrais, caractérisé du reste par une meilleure efficience d'utilisation.

Le Tableau I.3-2 indique les dates d'apports réalisés au cours de ces quatre campagnes ainsi que les doses apportées.

Saison	Apport 1	Apport 2	Apport 3	Apport 4
1999-2000	06/03 : 60	20/03 : 60	11/04 : 60	04/05 : 60
2000-2001	20/02 : 40	02/04 : 100	02/05 : 80	
2001-2002	04/03 : 50	28/03 : 50	10/04 : 40/70/100	06/05 : 0/30/60
2002-2003	11/03 : 40	26/03 : 60	23/04 : variable	12/05 : 40

Tableau I.3-2 : Dates des apports et quantités d'azote épandu (kg ha^{-1}) au cours des quatre campagnes. En 2002, les apports 3 et 4 ont été réalisés sur un carré latin en trois doses différentes. En 2003, la zone est de la parcelle a reçu un apport d'azote modulé, la zone ouest un apport constant de dose égale à la moyenne de celles de la zone est, soit 40 kg ha^{-1}.

Au cours des quatre saisons étudiées, des cartes de rendement et de teneurs en protéines ont été établies à la récolte et les teneurs en eau et en azote du sol au semis, en sortie d'hiver et à la récolte ont été mesurées.

I.3.5.2 Le jeu de données « essais doses d'azote »

Les jeux de données d'essais doses d'azote utilisés proviennent d'Arvalis – Institut du Végétal, de l'INRA de Grignon et de l'INRA de Laon (traitements azotés de Chambry). Une partie de ce jeu de données a servi à paramétrer STICS et le reste à le valider sur sa capacité à représenter la réponse du blé à la fertilisation azotée pour une large gamme de conditions pédo-climatiques caractéristiques des conditions du Nord de la France. Il a également servi à choisir une règle de préconisation de la fertilisation azotée.

I.4 Organisation de l'étude

La suite de ce travail est organisée en trois grands chapitres. Chacun d'eux comporte une analyse bibliographique concernant les points abordés.

Le Chapitre II concerne le volet sur l'évaluation de la simulation de la réponse du couvert à la fertilisation azotée par STICS et sur l'établissement d'une règle de décision concernant la fertilisation azotée prenant en compte des contraintes environnementales. Il utilisera le jeu de données « essais doses d'azote ». Le formalisme de STICS sera présenté à cette occasion.

Le Chapitre III, méthodologique, concerne la mise en œuvre des deux méthodes de spatialisation. La première, basée sur des observations et caractérisations pédologiques des parcelles, aboutira à la construction de bases de données permettant de renseigner le modèle STICS en tout point des deux parcelles. La seconde permettra, sur des jeux de données intégralement obtenus par simulation, de comparer trois méthodes d'assimilation et, plus généralement, de poser les bases de l'utilisation de techniques d'assimilation appliquées à STICS dans le cadre de l'agriculture de précision.

Le Chapitre IV utilisera le jeu de données de Chambry. Il permettra de mettre en œuvre dans le cas réel les deux approches développées au Chapitre III et de comparer leurs performances. L'assimilation sera utilisée seulement sur les saisons 1999-2000 et 2000-2001 : les saisons 2001-2002 et 2002-2003 serviront à évaluer la validité des estimations des variables d'entrée du modèle faites par assimilation pour une autre année climatique. Enfin, il décrira la mise en œuvre expérimentale de la méthode de préconisation de la fertilisation azotée développée dans le Chapitre II et analysera en première approche son intérêt.

Un chapitre de discussion générale et de conclusion permettra en définitive de tirer les principaux enseignements de cette étude.

II - Deuxième Chapitre Utilisation de STICS pour la préconisation de la fertilisation azotée

Une terre noire et grasse sous le soc qu'on enfonce et dont le sol est friable
(car c'est le résultat que nous cherchons à obtenir en labourant)
est presque toujours excellente pour les blés : en nulle autre plaine
tu ne verras plus de taureaux ramener à pas lents plus de chariots au logis.
(...) seulement n'aie point honte de saturer d'un gras fumier le sol aride,
ni de jeter une cendre immonde par les champs épuisés.

Virgile, Les Géorgiques

II.1 Introduction 51
II.2 Utilisation d'un modèle de culture pour préconiser des doses optimales d'azote: revue bibliographique 52
II.2.1 Critères agro-environnementaux 52
II.2.2 Méthodes utilisées pour déterminer des doses optimales par rapport à un critère donné 57
II.2.3 Evaluation d'un modèle 59
II.2.4 Démarche proposée 63
II.3 Présentation du modèle de culture et des données utilisés 65
II.3.1 Description du modèle STICS 65
II.3.2 Méthodologie de paramétrage de STICS 77
II.3.3 Données expérimentales 83
II.4 Paramétrage de STICS et évaluation de sa réponse à la fertilisation azotée 88
II.4.1 Résultats du paramétrage 88
II.4.2 Validation 92
II.4.3 Conclusion partielle 94
II.5 Choix d'un critère agro-environnemental 96
II.5.1 Proposition de différentes formes de critère 96
II.5.2 Application de ces critères pour deux niveaux de taxation ou de seuil ... 98
II.5.3 Application de ces critères pour une évolution continue des valeurs de taxation et de seuil 107
II.5.4 Conclusion partielle 113
II.6 Evaluation de l'aptitude de STICS à déterminer une stratégie de fertilisation azotée selon un critère agro-environnemental 114
II.6.1 Méthode d'évaluation utilisée 114
II.6.2 Validation des simulations réalisées avec des séries climatiques 117
II.6.3 Evaluation de la méthode pour deux niveaux de taxation ou de seuil ... 119
II.6.4 Evaluation de la méthode pour une évolution continue des valeurs de taxation et de seuil 121
II.7 Conclusion 125

II.1 Introduction

La première étape de la construction d'un outil de préconisation de la fertilisation azotée à l'échelle intra-parcellaire consiste à développer une méthode de détermination de la dose optimale d'azote basée sur l'utilisation d'un modèle de culture. Cette méthode n'étant pas spécifique à l'agriculture de précision, ce chapitre ne traitera donc pas de problèmes de spatialisation.

Une telle procédure implique d'abord de définir un critère de choix de la dose qui soit une fonction des différents facteurs que l'on veut optimiser. En effet, le mode de calcul de ce critère dépend du contexte dans lequel on se place : prise en compte ou non de facteurs économiques, environnementaux et de qualité. Dans la suite, nous l'appellerons « *critère agro-environnemental* », entendant par là qu'il est, plus ou moins exhaustivement et explicitement, une combinaison de ces différents facteurs. Chacun des termes impliqués dans le calcul du critère devra être bien prédit par le modèle pour différentes situations pédo-climatiques et différents niveaux de fertilisation azotée. Il conviendra donc d'estimer la qualité de la simulation de ces termes grâce à un jeu de données alliant une variabilité importante des types de sols et des doses d'azote.

Une fois le critère établi et le modèle évalué par rapport à sa prédiction, il faudra développer une méthode d'optimisation de la dose ou du scénario technique au moment du choix des apports par rapport à ce critère, en prenant en compte l'incertitude sur le climat à venir. Enfin, l'aptitude du modèle à effectuer des bonnes préconisations de fertilisation azotée devra être évaluée selon deux aspects : (i) comparer les choix effectués par le modèle par rapport au meilleur des choix possibles déterminé *a posteriori*, c'est-à-dire d'après les données expérimentales et (ii) évaluer les conséquences du choix réalisé par le modèle sur le critère retenu. Une mauvaise recommandation peut par exemple avoir de faibles conséquences si la réponse du critère à la dose d'azote est faible. Nous avons également choisi d'évaluer les résultats obtenus avec STICS par rapport à ceux obtenus avec une méthode de référence en terme de préconisation de la fertilisation azotée et largement utilisée en France, à savoir AZOBIL (Machet *et al.*, 1990).

II.2 Utilisation d'un modèle de culture pour préconiser des doses optimales d'azote: revue bibliographique

De nombreux outils informatiques de préconisation de la fertilisation azotée basés sur l'utilisation de modèles agronomiques existent déjà, surtout à l'échelle de la parcelle entière. Certains d'entre eux sont inclus dans des outils plus généraux de préconisation d'itinéraires techniques (Engel, 1997 ; Shaffer & Brodahl, 1998 ; Booltink *et al.*, 2001 ; Loyce *et al.*, 2002 ; Jones *et al.*, 2003).

Cette partie nous permettra de :

- passer en revue différentes formes de critères agro-environnementaux, c'est-à-dire la variable qui permet d'optimiser la dose d'azote qui synthétise des contraintes agronomiques et environnementales ;

- décrire les façons d'utiliser un modèle de culture pour préconiser les doses optimales d'azote utilisant ces critères ;

- d'introduire les modalités utilisées pour évaluer le modèle ;

- présenter la démarche utilisée dans le reste de ce chapitre.

II.2.1 Critères agro-environnementaux

La définition d'un itinéraire technique de fertilisation azotée vise à atteindre différents objectifs agronomiques et environnemetaux pouvant être contradictoires et relève d'un problème d'optimisation multicritère. Les différentes composantes à prendre en compte sont le rendement en grains (G), la qualité des grains *via* la teneur en protéines (P) qui doivent être maximisés, et l'impact environnemental (E) qui doit être minimisé. Ces trois variables de sortie du modèle, de type $\hat{Y}^S = f(X;\theta)$, doivent être combinées en un critère commun, que l'on peut nommer *fonction objectif*, notée J et dont l'utilisation correspond au critère agro-environnemental. Une fonction objectif simulée sera notée $\hat{J}$. Cette fonction objectif doit être optimisée par rapport à l'itinéraire technique relatif à la dose d'engrais azoté, notée N. La dose optimale sera notée N^* dans le cas expérimental et $\hat{N}^*$ dans le cas de l'utilisation du modèle.

N^* est en réalité dépendant de X et $\hat{N}^*$ de X et de θ. Pour plus de commodité, nous noterons la plupart du temps N^* au lieu de $N^*(X)$ et $\hat{N}^*$ au lieu de $\hat{N}^*(X;\theta)$.

Différentes fonctions objectifs sont présentées dans les parties qui suivent. La première fait exception au schéma présenté ci-dessus car elle est basée sur l'utilisation de variables d'état. Pour les autres fonctions objectifs, les composantes à prendre en compte dans un critère agro-environnemental (G, P et E) ne sont pas forcément toutes explicitement prises en compte.

II.2.1.1 Calcul d'une balance azotée dynamique

Engel (1997), Jeuffroy & Recous (1999) et Booltink *et al.* (2001) proposent des méthodes voisines, toutes deux fondées sur l'utilisation de modèles dynamiques : elles n'utilisent pas des variables de type G, P et E, caractéristiques de la récolte, mais se basent sur une approche assez similaire à celle de la balance azotée (Stanford, 1973), impliquant des variables d'état obtenues en cours de culture, de type $\hat{Y}(t)= f(X;\theta;t)$. Le critère proposé par ces auteurs est la différence entre les besoins de la plante et la quantité d'azote que le sol est susceptible de fournir. La façon dont la dose d'azote est calculée sera décrite dans la partie II.2.2.

Dans cette méthode qui vise à ajuster en temps réel les fournitures aux besoins, on optimise implicitement le rendement car on calcule des doses qui permettent de minimiser les carences azotées. L'aspect environnemental est pris indirectement en compte par le fait que les apports sont les plus proches possibles des besoins, ce qui devrait conduire à minimiser les reliquats à la récolte. Cette méthode s'apparente beaucoup, dans le principe, aux « méthodes indicateurs » (Chapitre I) puisque, dans un cas comme dans l'autre, on calcule des besoins en azote au moment des apports. Elle présente de plus l'avantage de préconiser les dates d'apport.

II.2.1.2 Détermination de la dose optimale selon la courbe de réponse du rendement à la dose d'azote

Cette méthode calcule une dose d'azote optimale par ajustement de modèles statiques à des données expérimentales, selon le principe utilisé notamment par Laurent (2000) et Makowski *et al.* (2001). La Figure II.2-1-a montre une forme possible de courbe de réponse du rendement en grain (G) à la dose d'azote : la dose optimale N^* correspond à un des paramètres de celle-ci, et peut être définie comme la dose minimale qui maximise le

rendement. L'aspect environnemental n'est pas explicitement pris en considération ici, mais le fait de prendre la « dose minimale » correspond aussi implicitement à des préoccupations d'ordre environnemental et économique.

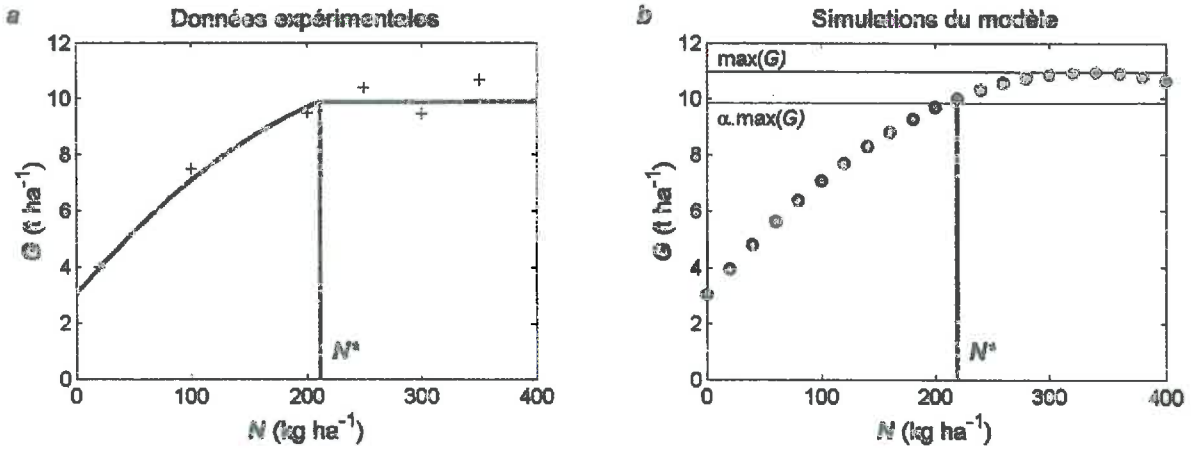


Figure II.2-1 : a- Exemple de forme d'ajustement d'un modèle de réponse du rendement sur des données expérimentales et détermination de la dose d'azote optimale ; b- Réponse du rendement simulé par un modèle de culture à la dose d'azote et proposition d'un mode de détermination de la dose optimale.

On pourrait généraliser cette méthode au cas où la courbe de réponse du rendement à la dose d'azote serait le résultat de simulations de modèles de culture réalisées avec un pas fin de doses d'azote (Figure II.2-1-b). Comme avec un modèle de culture $\hat{G}$ n'atteint pas forcément un plateau strict, la dose optimale $\hat{N}^*$ peut être définie comme la dose minimale permettant d'atteindre une certaine proportion α du rendement maximal. On aurait alors :

$$\hat{J}(N)=\hat{G}(N) \qquad \text{Équation II.2-1}$$

$$\hat{N}^*=\min[N \text{ tel que } \hat{J}(N)\geq\alpha.\hat{J}(N)] \qquad \text{Équation II.2-2}$$

Les deux types de critères décrits dans la partie précédente et celle-ci ne permettent pas de prendre en compte des facteurs de qualité comme la teneur en protéines *P*.

II.2.1.3 Création d'une variable synthétique de *G*, *P* et *E* par pondération

Dans un cadre plus complexe d'évaluation de l'ensemble de l'itinéraire technique, Loyce *et al.* (2002) utilisent une analyse multicritère pour classer des schémas d'opérations culturales. Pour chacun d'entre eux, l'utilisateur choisit des classes de valeurs acceptables qui définissent des intervalles de rejet et d'acceptation sous la forme d'une fonction de degré d'appartenance aux classes « acceptée » et « rejetée », comprise entre 0 et 1. On évalue

ensuite un critère synthétique de rejet et d'acceptation combinant, selon des poids également choisis par l'utilisateur, les critères calculés pour chaque variable. Avec notre exemple sur G, P et E, ceci pourrait donner ce qui est illustré dans la Figure II.2-2 : le critère d'acceptation pour le rendement, $J_G(N)$, est une fonction 'plateau-linéaire croissante-plateau' du rendement, comprise entre 0 et 1 et caractérisée par les valeurs g_1 et g_2. Si on ne prenait en compte que ce critère, ceci signifierait que toutes les doses N telles que $J_G(N)=1$ seraient acceptées et celles telles que $J_G(N)=0$ seraient refusées. De même, $J_P(N)$ est une fonction 'plateau-linéaire croissante-plateau' de P et $J_E(N)$ une fonction 'plateau-linéaire décroissante-plateau' de E.

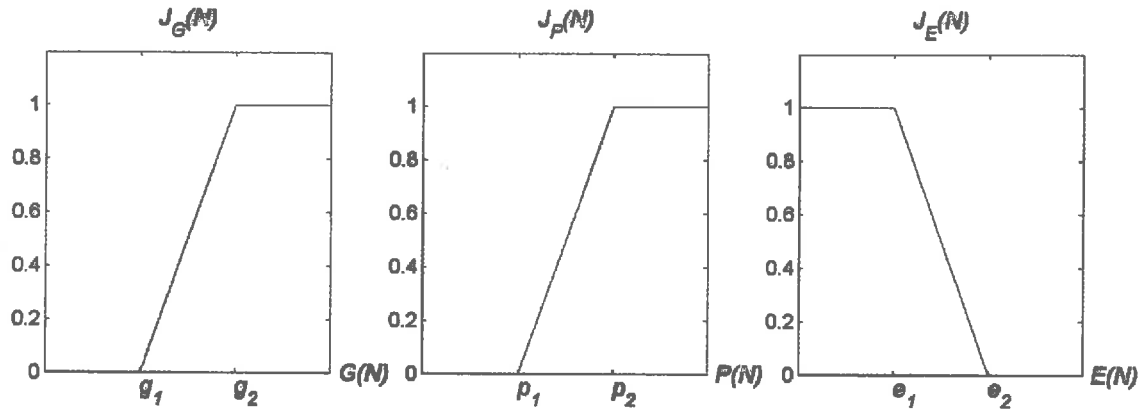


Figure II.2-2 : Fonction de degré d'appartenance à la catégorie « acceptée » pour les trois variables du critère agro-environnemental dans une approche d'analyse multicritère.

Le critère global, ou fonction objectif, se calculerait ensuite de la façon suivante :

$$J(N)=w_G.J_G(N)+w_P.J_P(N)+w_E.J_E(N) \qquad \text{Équation II.2-3}$$

où w_G, w_P et w_E sont des facteurs de pondération qui sont, au même titre que g_1, g_2, p_1, p_2, e_1 et e_2, des valeurs à fixer par l'utilisateur. Cette fonction objectif traduit l'adéquation des résultats induits par une dose à un ensemble de critères. La façon de déterminer $\hat{N}^*$ serait la suivante :

$$\hat{N}^*=N \text{ tel que } \hat{J}(N)=\max[\hat{J}(N)] \qquad \text{Équation II.2-4}$$

On peut aussi, ce qui est intéressant, rajouter une contrainte pour refuser systématiquement des valeurs de N conduisant à des valeurs rédhibitoires d'un des critères d'acceptation.

Cette approche aurait pu être utilisée dans notre cas, mais elle semble plutôt adaptée à des situations plus complexes telles que celle traitée par Loyce *et al.* (2002), impliquant des aspects très différents de l'itinéraire technique dont les fonctions d'acceptation ne sont pas

toujours continues. En outre, se pose le problème du choix, subjectif, des valeurs des différents poids w et bornes g, p et e.

II.2.1.4 Détermination de la dose optimale selon la courbe de réponse de la marge brute à la dose d'azote

Dans le cas où les critères peuvent s'exprimer de façon continue, comme G et P, une façon réaliste de les combiner est de les utiliser pour calculer la marge brute. La fonction objectif s'écrit alors :

$$J(N)=G(N)\times f[P(N)]-a.N \qquad \text{Équation II.2-5}$$

où $J(N)$ représente la marge brute (€ ha^{-1}) obtenue avec la dose N (kg ha^{-1}), $f[P(N)]$ est la fonction qui donne le prix du blé (€ kg^{-1}) en fonction de sa teneur en protéines et a est le prix de l'engrais azoté (€ kg^{-1}). Pour définir f, on peut utiliser le barème appliqué par les coopératives agricoles, qui présente en général la forme d'une fonction 'plateau-linéaire croissante-plateau'. On peut d'ores et déjà remarquer que ce formalisme donne peu de poids à P.

La dose optimale est alors celle qui maximise la marge brute (par exemple Paz *et al.*, 1999, Makowski *et al.*, 2001) :

$$\hat{N}^*=N \text{ tel que } \hat{J}(N)=\max[\hat{J}(N)] \qquad \text{Équation II.2-6}$$

Là encore, la préoccupation environnementale n'est qu'indirectement prise en considération. Le fait d'inclure le prix de l'engrais azoté dans le calcul de la fonction objectif a tendance à diminuer la dose retenue. Mais avec les coûts actuels des engrais azotés, relativement faibles, cette contrainte reste peu importante et équivalente à celle qui est le fait du choix de la dose minimale qui maximise le rendement (*cf.* II.2.1.2).

II.2.1.5 Ajout d'une contrainte environnementale

Une façon de prendre en compte explicitement une contrainte environnementale est d'inclure la variable E dans le calcul de la marge brute. Makowski *et al.* (2001) proposent de fonder le calcul de E sur les reliquats d'azote à la récolte : $E(N)=R(N)$. La fonction objectif s'écrit alors :

$$J(N)=G(N)\times f[P(N)]-a.N-g[R(N)] \qquad \text{Équation II.2-7}$$

où g représente une fonction de pénalisation. Comme aucune mesure n'existe actuellement au moins en France, on peut proposer plusieurs formes pour la fonction g. Ces auteurs suggèrent la possibilité de considérer cette fonction nulle en deçà d'un seuil de reliquat azoté de 35 kg ha^{-1} (correspondant approximativement à la norme européenne de potabilité de l'eau de 50 mg L^{-1}) et égale, au-delà, au coût de mise en place d'une culture intermédiaire, estimé à 38 € ha^{-1}.

Il existe un grand nombre de variantes possibles pour décrire la fonction g. elle pourrait dépendre d'autres variables que R. On peut également citer les termes d'un projet de loi soumis au parlement français mais ajourné au cours de l'été 2002 (Laurent, Machet, comm. pers.). Ce projet de loi prévoyait de calculer, à l'échelle de l'exploitation agricole, un bilan apparent d'azote dans le sol (entrées moins sorties). C'est ce bilan qui, rapporté à l'échelle de la parcelle, pourrait correspondre à notre variable E et pourrait être soumis à une taxation g. Il s'agit donc d'une variante sur la forme de la fonction g et sur la nature de E. Ce bilan d'azote, noté B, pourrait se calculer à l'échelle de la parcelle ou de la placette de la façon suivante :

$$B(N)=N\times-10.G(N).P(N)/5.7 \qquad \text{Équation II.2-8}$$

On considère donc ici les sorties uniquement sous la forme de l'azote exporté par les grains. D'après Grundy *et al.* (1996), le facteur permettant de passer de la teneur en azote des grains à la teneur en protéines est de 5.7.

La façon de choisir $\hat{N}^*$ serait la même que celle décrite par l'Equation III.2-6.

II.2.2 Méthodes utilisées pour déterminer des doses optimales par rapport à un critère donné

Selon le critère choisi pour calculer une dose optimale d'azote, la méthode mise en œuvre peut ne pas être la même. Lorsque le critère est construit sur des variables d'état intermédiaire (calcul des besoins en azote de la culture, *cf.* II.2.1.1), on peut raisonner à un pas de temps relativement bref, d'une date d'apport à l'autre par exemple. Par contre, lorsque les fonctions objectifs sont construites sur les variables d'intérêt final (II.2.1.2 à II.2.1.5), on est obligé de mener la simulation jusqu'à la récolte. Dans un cas comme dans l'autre, l'incertitude climatique doit être prise en compte par l'utilisation de scénarios climatiques.

II.2.2.1 Méthode de calcul d'une balance azotée dynamique

Le paragraphe qui suit est un résumé des méthodes que proposent Engel (1997), Jeuffroy & Recous (1999) et Booltink *et al.* (2001), qui sont assez similaires.

Entre deux dates d'apports, on simule de façon prédictive en utilisant des séries climatiques (une vingtaine) la quantité d'azote absorbée par la culture d'une part et la quantité d'azote minéral dans la zone racinaire d'autre part. Lorsque cette dernière devient inférieure à un seuil qui dépend de la vitesse d'absorption, on apporte la quantité d'azote correspondant à la consommation prévue pour les quatre semaines suivantes. Dans le cas du dernier apport, la simulation va jusqu'à la récolte. A chaque apport, on prend en compte l'effet du climat réel, ce qui implique de réactualiser les simulations.

Cette méthode permet de se dispenser (i) d'utiliser de longues séries climatiques (*i.e.* jusqu'à la récolte) qui sont sources d'erreur au final et (ii) de déterminer à l'avance plusieurs itinéraires techniques à comparer. Néanmoins, elle ne considère que le système sol plante et ne prend pas en compte les contraintes économiques et environnementales de façon explicite. Elle pourrait cependant être utile pour déterminer la date des apports.

II.2.2.2 Méthode de calcul d'une fonction objectif définie à la récolte

Cette méthode s'applique aux critères évoqués dans les parties II.2.1.2 à II.2.1.5. Comme il s'agit d'obtenir des variables à la récolte, on peut utiliser soit des modèles statiques (Makowski *et al.*, 2001 ; Loyce *et al.*, 2002) soit des modèles dynamiques (Paz *et al.*, 1999), l'utilisation de ces derniers permettant de prendre en compte les facteurs pédo-climatiques. On compare les variables simulées à la récolte pour différents scénarios de fertilisation azotée et de climat futur. On procède donc à des simulations qui croisent :

- n scénarios climatiques (pour les modèles qui en ont besoin) à partir de la date de décision ;

- m scénarios techniques variant par la dose d'azote et éventuellement les dates d'apport.

A titre d'exemple, Paz *et al.* (1999) croisent n=22 climats et m=21 doses d'azote. Les variables utilisées pour le calcul de la fonction objectif peuvent être analysées sous forme de moyenne climatique ou sous forme fréquentielle de façon à permettre une analyse de risque.

Cette méthode implique d'utiliser des séries climatiques allant de la date de décision à la récolte et de prédéfinir des itinéraires techniques. Mais son utilisation permet le calcul de critères plus complexes, intégrés sur l'ensemble du cycle de la culture et prenant en compte des aspects économiques et environnementaux.

Le choix s'est assez rapidement porté sur des fonctions objectifs définies à la récolte car elles permettent de quantifier l'impact de la fertilisation sur l'environnement. Par la suite, nous ne parlerons donc plus des critères basés sur le suivi des stress azotés en continu.

II.2.3 Evaluation d'un modèle

Dans le cadre que nous nous proposons, l'évaluation du modèle de culture doit comporter deux phases :

- l'évaluation de la capacité du modèle à décrire l'évolution des variables d'état et à prévoir les variables d'intérêt final ;

- l'évaluation de sa capacité à faire des préconisations de fertilisation azotée selon un critère agro-environnemental et des conséquences de ces préconisations si elles étaient appliquées.

Dans cette partie, nous définirons la RMSE utilisée pour le premier aspect et nous expliquerons une méthode proposée par Wallach (2002) pour traiter du second point.

II.2.3.1 Définition de la RMSE

Pour évaluer la qualité de la réponse de STICS à la fertilisation azotée et à la variabilité des caractéristiques du sol, nous utiliserons essentiellement le critère statistique nommé racine carrée de l'erreur quadratique moyenne (« Root Mean Square Error », RMSE) et défini de la façon suivante :

$$RMSE=\sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(Y_i-\hat{Y}_i\right)^2} \qquad \text{Équation II.2-9}$$

où Y_i représente la $i^{ème}$ valeur observée, $\hat{Y}_i$ la $i^{ème}$ valeur simulée et I le nombre d'observations.

En outre, afin de quantifier la part de l'erreur mesurée par la RMSE qui est due à un biais du modèle et celle qui est due à la dispersion, nous utiliserons la RMSEs (s pour « systematic ») et la RMSEu (u pour « unsystematic) comme proposé par Willmott (1981) et

définies comme suit. Ces notions se basent sur la régression de $\hat{Y}_i$ en fonction de Y_i. Le résultat de cette régression est une relation linéaire : $\hat{Y}_i^* = a + b.Y_i$ où $\hat{Y}_i^*$ est la valeur de $\hat{Y}_i$ prédite par la régression linéaire. On a :

$$RMSEs = \sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(Y_i - \hat{Y}_i^*\right)^2} \qquad \text{Équation II.2-10}$$

$$RMSEu = \sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(\hat{Y}_i - \hat{Y}_i^*\right)^2} \qquad \text{Équation II.2-11}$$

La RMSEs mesure l'écart entre la droite de régression et la première bissectrice alors que la RMSEu mesure la dispersion autour de cette droite de régression des valeurs simulées par le modèle, $\hat{Y}_i$ (Figure II.2-3).

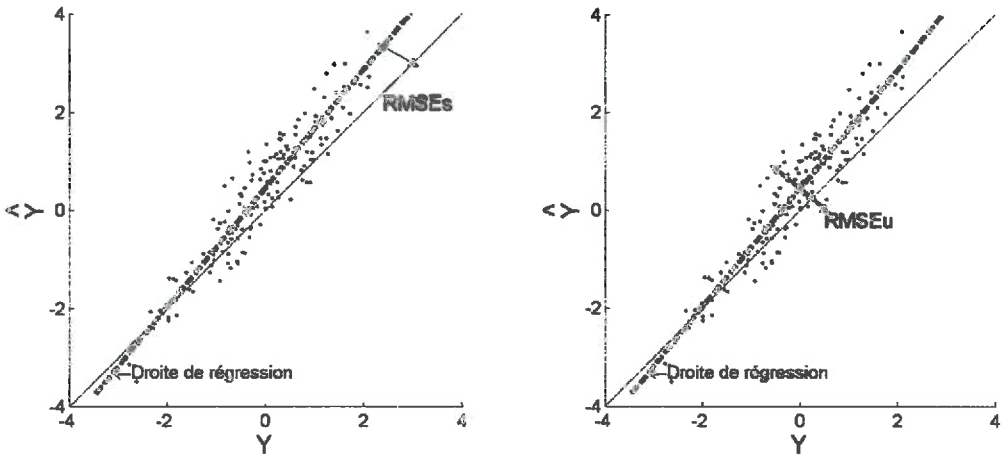


Figure II.2-3 : Décomposition de la RMSE (Root Mean Square Error) en RMSE systématique (RMSEs) et non systématique (RMSEu). a) La RMSEs mesure le biais introduit par le modèle ; b) la RMSEu mesure la dispersion autour de la droite de régression entre valeurs simulées et observées. Y : valeurs observées ; $\hat{Y}$: valeurs simulées par le modèle.

On utilisera également les RRMSE, RRMSEs, RRMSEu qui sont respectivement les RMSE, RMSEs et RMSEu divisée par la moyenne des observations.

II.2.3.2 Evaluation de la capacité d'un modèle à effectuer des préconisations

Un modèle prédisant relativement bien la fonction objectif peut conduire à des préconisations moins bonnes qu'un autre modèle jugé moins performant en terme de qualité prédictive. La Figure II.2-4 représente sur un exemple théorique l'évolution réelle, en trait plein, de la fonction objectif J, d'après les observations. On constate que J est maximale pour

une valeur de N égale à N_2 et vaut alors J_2. Le modèle A (traits pointillés) prévoit fort mal la valeur de J mais prédit parfaitement que la marge J atteint sa valeur maximale pour $N=N_2$.

Dans cette situation, la préconisation de N par ce modèle serait idéale. En revanche, le modèle B (traits fins) prévoit relativement mieux les valeurs de J en fonction de N mais donne N_1 comme valeur de N maximisant J au lieu de N_2. L'apport de la dose N_1 conduit en réalité à une valeur de la fonction objectif égale à J_1, très inférieure à J_2. Il y a deux aspects à prendre en compte : d'une part l'erreur commise par le modèle B lors de la préconisation et d'autre part la conséquence qu'entraîne cette erreur. La première est caractérisée par la différence entre N_1 et N_2 ; la seconde par la différence entre J_1 et J_2.

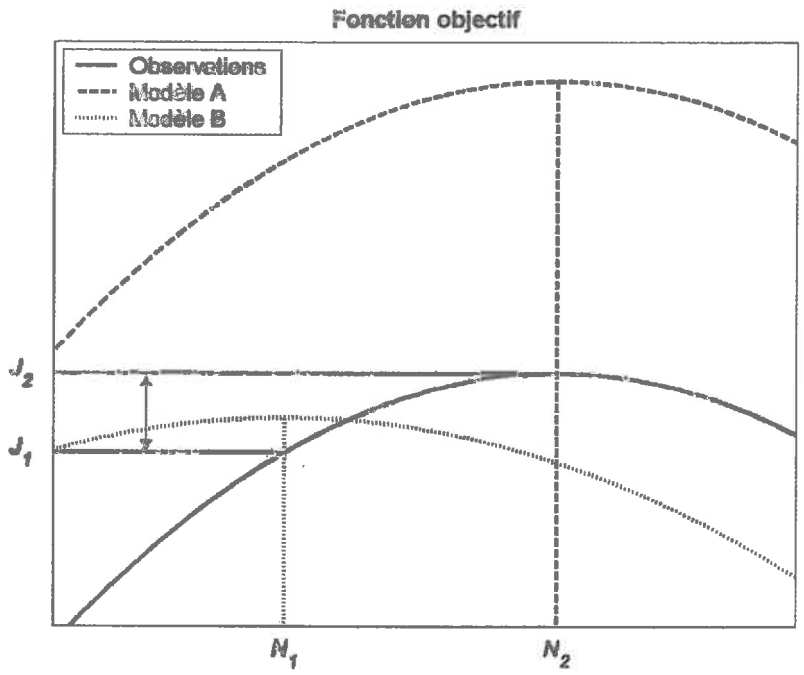


Figure II.2-4 : Courbes de réponse théoriques illustrant la différence entre qualités de prédiction et de préconisation d'un modèle. D'après Wallach (2002).

Sur la base de cet exemple, Wallach (2002) propose d'évaluer la qualité de préconisation d'un modèle de la façon suivante. Sachant que les résultats du modèle, tant en terme de prédiction que de préconisation, dépendent des paramètres $\hat{\theta}$ qui sont utilisés dans celui-ci, le critère suivant peut être utilisé pour juger de la qualité de préconisation d'un modèle :

$$C(\hat{\theta})=E[J(\hat{N}^*(X;\hat{\theta}))|\hat{\theta}] \qquad \text{Équation II.2-12}$$

où $C(\hat{\theta})$ représente le critère d'évaluation, dépendant des paramètres utilisés par le modèle, J est la fonction objectif réelle, $\hat{N}^*(X;\hat{\theta})$ représente la préconisation effectuée par le modèle en fonction des paramètres $\hat{\theta}$ et des conditions représentées par X, et E représente l'espérance. L'idée sous-jacente est que, le but étant de maximiser la fonction objectif J, il faut évaluer la qualité des préconisations du modèle d'après la valeur qu'elle aurait prise en appliquant les décisions du modèle. Dans l'exemple ci-dessus, $C(\hat{\theta})$ vaut J_2 pour le modèle A et J_1 pour le modèle B. On peut ainsi comparer deux modèles entre eux ou un modèle par rapport à la référence expérimentale : dans le cas illustré, c'est la préconisation du modèle A qui est la meilleure.

Néanmoins, en règle générale, on ne connaît pas la fonction objectif réelle de façon continue. Si d'aventure, ce qui est fort probable, le modèle préconise une dose d'azote qui n'a pas été réellement expérimentée, on ne peut pas déterminer la valeur vraie de la fonction objectif pour cette dose, c'est-à-dire $J(\hat{N}^*(X;\hat{\theta}))$; par exemple, si N_1 n'avait pas été appliquée, on ne pourrait pas connaître J_1. Face à cette situation, deux solutions peuvent s'appliquer :

- estimer $J(\hat{N}^*(X;\hat{\theta}))$ en interpolant les valeurs des mesures entre les doses expérimentées comme proposé par Antoniadou et Wallach (2000). Ceci n'est possible que lorsqu'on considère des décisions simples, une seule dose d'azote par exemple ;

- ne considérer que les décisions réellement appliquées.

Cette dernière solution est illustrée par la Figure II.2-5 qui compare les valeurs observées et simulées de la fonction objectif dans trois situations. Dans la première, le modèle préconise la dose 300 kg ha^{-1} qui correspond au maximum de la fonction objectif simulée. La valeur réelle de la fonction objectif pour cette décision est 720. Dans la situation 2, le modèle préconise la dose 250 qui conduit à une valeur réelle de fonction objectif 775. Pour la situation 3, on obtient le couple de valeurs (200;950). La fonction objectif simulée vaut en moyenne : $\overline{C}(\hat{\theta}) = 815$. Cette valeur est à comparer à celle obtenue par un autre modèle ou à la valeur maximale du critère qu'il est possible d'atteindre, c'est-à-dire celle qui est calculée à partir des observations, ici : $\hat{C}_{max} = (750+900+950)/3 = 867$. On peut ainsi parler de « perte » engendrée par l'utilisation du modèle par rapport à l'utilisation d'un autre modèle ou par rapport à la situation idéale, ici 52 (Makowski et Wallach, 2001).

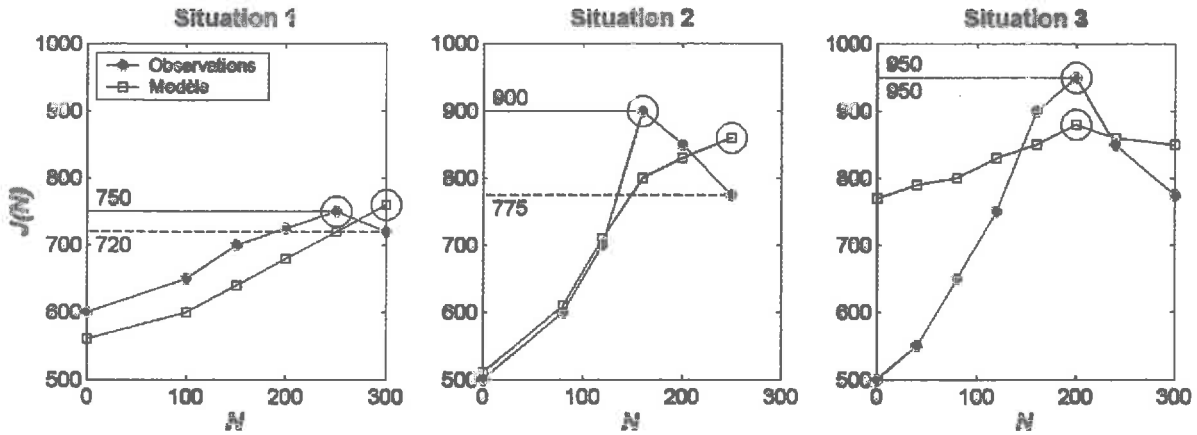


Figure II.2-5 : Exemples illustrant le calcul du critère d'évaluation de la capacité d'un modèle à réaliser des préconisations. Dans la situation 1, l'utilisation du modèle conduit à une perte de fonction objectif *J* égale à 30. La perte est de 125 dans la situation 2 et de 0 dans la situation 3.

Cet exemple permet en outre de faire quelques constatations simples et intuitives sur le comportement de ce critère :

- les situations 1 et 2 montrent que la forme de la réponse de la fonction objectif à la variable que l'on veut optimiser influence les valeurs relatives de $\overline{C}(\hat{\theta})$ et de $\hat{C}_{max}$. Si celle-ci est assez plate comme dans la situation 1, cela aura peu de conséquences sur l'écart entre $\overline{C}(\hat{\theta})$ et de $\hat{C}_{max}$; si la réponse de la fonction objectif est forte (situation 2), l'écart sera beaucoup plus marqué ;

- les situations 2 et 3 montrent que la qualité de prédiction du modèle n'est pas forcément liée à la qualité de la préconisation : dans la situation 2, le modèle réalise de très bonnes simulations pour les faibles doses d'azote, mais cela n'a aucune influence sur sa capacité à réaliser des préconisations. *A contrario*, dans la situation 3 le modèle simule mal la courbe de réponse à l'azote mais prévoit bien la dose optimale.

II.2.4 Démarche proposée

L'étude bibliographique précédente nous amène à préciser les étapes nécessaires pour évaluer la qualité de préconisation de fertilisation azotée réalisée par un modèle de culture tel que STICS. Les six étapes sont résumées dans la Figure II.2-6.

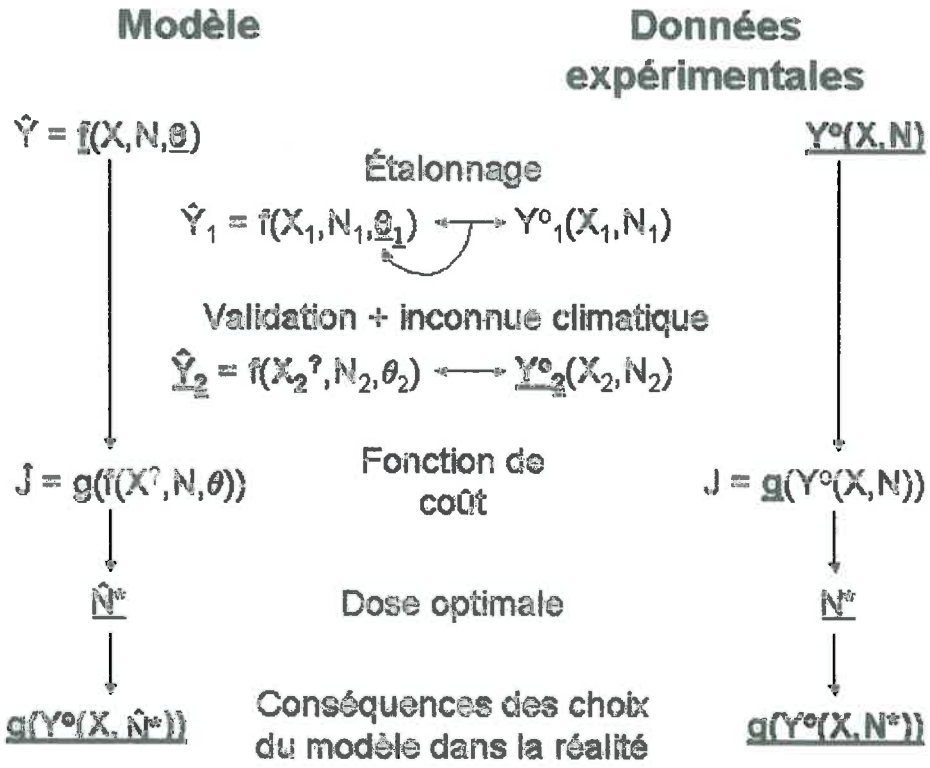


Figure II.2-6 : Démarche proposée : les différentes étapes sont présentées en ligne ; en gras et souligné sont indiqués les aspects concernés plus particulièrement par les étapes correspondantes.

1. Présentation des données expérimentales, représentées par : X (conditions pédologiques, climatiques, techniques) ; N (doses d'azote) ; Y^o (variables d'intérêt final observées, G, P et E, avec $E=R$ ou $E=B$).

2. Description du modèle, f, notamment par rapport à sa façon de décrire l'influence des propriétés du sol et des apports azotés sur $\hat{Y}$. Choix des paramètres θ du modèle qui sont importants ou susceptibles d'être paramétrés.

3. Paramétrage du modèle sur une partie du jeu de données et validation sur l'autre partie (confrontation $\hat{Y}$ / Y^o). La validation se fera soit *a posteriori* soit en ignorant le climat réel à partir de la deuxième date d'apport, en utilisant des séries climatiques (notées $X^?$).

4. Choix de la fonction objectif, $J=g(Y^o(X,N))$, en utilisant les données expérimentales et de la variable contenue dans E impliquée dans le calcul de J.

5. Détermination de la dose optimale d'azote observée (N^*) et de la dose optimale simulée ($\hat{N}^*$).

6. Evaluation des conséquences des choix du modèle par rapport à ceux d'une méthode de référence, AZOBIL (Machet *et al.*, 1990).

II.3 Présentation du modèle de culture et des données utilisés

Il va dans le marais qui a nom Styx,
le sinistre ruisseau, quand il arrive
au pied des affreuses berges grises.
Et moi qui regardais très fixement,
je vis des gens boueux dans le marais,
tous nus, à l'aspect meurtri.
Dante, La Divine Comédie, L'enfer, VII, 106-111

Cette partie présente le modèle STICS en insistant sur les aspects de la réponse du rendement, de la teneur en protéines et du reliquat d'azote minéral à la récolte tant à la fertilisation azotée qu'aux propriétés du sol. Les paramètres essentiels dans la description de ces phénomènes et qui feront l'objet d'un paramétrage seront dégagés. Les données présentées par la suite serviront à la fois à réaliser ce paramétrage, à déterminer des critères agro-environnementaux pertinents et à évaluer l'aptitude du modèle à préconiser des doses d'azote.

II.3.1 Description du modèle STICS

Le modèle de culture STICS (Simulateur mulTIdisciplinaire pour les Cultures Standard) a été initié en 1996 ; une description exhaustive des principes généraux de fonctionnement de sa version 3 a été faite par Brisson *et al.* (1998). Une autre publication présente une comparaison des résultats des versions 3 et 4 (Brisson *et al.* 2002). Une analyse de sensibilité par module a été effectuée pour la version 3 par Ruget *et al.* (2002). Les caractéristiques de la version 5, sortie au cours de l'été 2002, ont été présentées par Brisson *et al.* en 2003. Les travaux présentés par la suite ont été réalisés avec la version 5.

Les sous-parties qui suivent présentent d'abord les grands traits du fonctionnement du modèle, puis détaillent davantage son aspect modulaire et les relations entre les différentes variables qui nous intéressent, de façon à dégager des paramètres importants du point de vue du paramétrage. Une description plus exhaustive des variables d'entrées sera ensuite réalisée. Enfin, on donnera des résultats d'évaluation du modèle STICS obtenus dans le cadre d'autres études de façon à servir de base de comparaison avec nos propres résultats de validation.

II.3.1.1 Brève présentation

STICS fonctionne à un pas de temps journalier. Il simule les transferts d'eau, de carbone et d'azote, de façon mono-dimensionnelle, dans un système sol-plante limité par l'atmosphère d'une part et la frontière sol/sous-sol d'autre part. Les principes ayant servi à sa construction sont la robustesse (aptitude à simuler des conditions pédo-climatiques variées), la généricité (adaptabilité à un grand nombre de cultures) et la modularité (possibilité de rajouter la description de mécanismes non décrits initialement).

Les entrées concernent (Figure II.3-1) :

- les variables climatiques : températures minimales et maximales, pluviométrie, rayonnement et évapotranspiration potentielle au pas de temps journalier ;
- les propriétés permanentes du sol, sur cinq horizons : propriétés hydriques et chimiques ;
- les caractéristiques de la culture : développement, croissance et sensibilité aux stress ;
- l'itinéraire cultural : dates d'intervention, apports d'eau et d'azote, semis... ;
- l'état initial du sol : teneurs en eau et en azote minéral ;

Les sorties concernent principalement :

- le rendement en grains et leur qualité (teneur en azote des grains) ;
- des données environnementales : drainage de l'eau et lixiviation des nitrates ;
- l'état final du sol : teneurs en eau et en azote minéral.

L'entité élémentaire de simulation dans STICS correspond à une zone homogène par rapport à tous les types d'entrées (Figure II.3-1). Dans ce chapitre, elle correspond à la placette (1 m^2 environ) qui sert de support à l'essai azote. Le modèle assimile la culture à une couche décrite par sa biomasse totale (*MS*), son indice de surface foliaire (*LAI*), sa quantité d'azote absorbée (*QN*) et sa biomasse récoltée (*G*). Le sol est schématisé par des couches d'épaisseur variable et de propriétés spécifiques. L'interaction entre le sol et la culture se fait

par le biais des racines, décrites en terme de profondeur maximale et de densité (Brisson *et al.*, 2003).

Le modèle STICS s'articule autour de trois grandes fonctions (Ruget *et al.* 2002) :

- le calcul des stades de développement d'après un indice vernalo-photothermique ;
- la croissance potentielle de la culture pour chaque stade végétatif, déterminée par les variables climatiques et les caractéristiques de la culture ;
- l'effet des stress hydriques et azotés sur la croissance potentielle qui requiert la simulation de la croissance des racines.

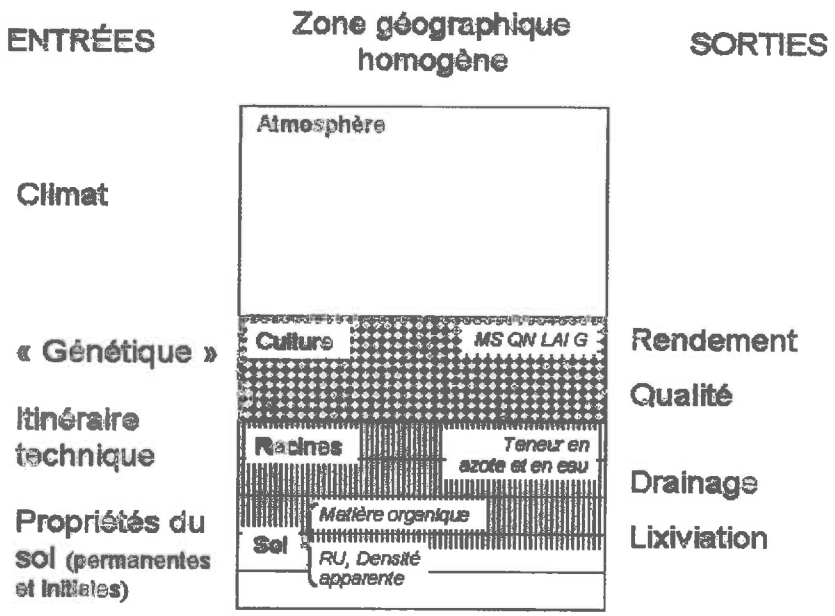


Figure II.3-1 : Système modélisé par STICS et ses principales entrées et sorties.

STICS est organisé en neuf modules qui sont fortement interactifs. Ils communiquent entre eux par diverses variables d'état. Les paragraphes suivants présenteront (i) cette organisation modulaire, avec les entrées et sorties de chaque module ; (ii) les principales relations existant entre variables d'entrées, variables d'état et variables de sortie ; (iii) les principaux paramètres à calibrer.

II.3.1.2 Organisation modulaire de STICS : les principaux processus simulés

Cette présentation, plus complète que le compendieux résumé précédent, est volontairement non exhaustive et simplificatrice, se bornant à ne donner que les éléments essentiels à la compréhension des travaux présentés dans la suite de ce document. En particulier, on ne décrit que les formalismes correspondant aux options de simulation choisies (*e.g.* pour les racines, l'option densité vraie). La Figure II.3-2 représente les principales entrées et sorties des modules de STICS (Brisson *et al.*, 2003). Dans ce qui suit, les paramètres essentiels sont écrits en caractères minuscules italiques et seront synthétisés dans la partie II.3.1.4, les variables d'état indiquées en caractères majuscules italiques et les grands facteurs déterminant les processus signalés en caractères minuscules gras.

- Un indice de développement, *ULAI(t)*, est calculé par le **module Phénologie** à partir des données **climatiques** et des durées de stades de développement caractéristiques de **l'espèce et de la variété**. Les paramètres de stades de développement sont *stlevamf*, correspondant à la durée entre la levée et l'accélération maximale de l'accroissement de *LAI* et *stamflax*, durée entre le stade précédent et celui marquant la fin de la croissance foliaire.

- La germination est une fonction du temps et de la température ; le principal paramètre qui la gouverne est *spltger*, cumul d'unités de développement permettant la germination, caractéristique de **l'espèce**.

- Le **module Croissance aérienne** calcule d'abord *LAI* en fonction du développement. L'accroissement potentiel journalier de *LAI*, $\Delta LAI(t)$, est calculé selon la formule :

$$\Delta LAI(t)=\frac{dlai_{max}}{1+\exp[pentlai_{max}.(vlai_{max}-ULAI(t))]}\times(TCULT(t)-tc_{min}) \qquad \text{Équation II.3-1}$$

Les paramètres $dlai_{max}$, $pentlai_{max}$ et $vlai_{max}$ sont spécifiques de **l'espèce**. L'effet de la température de la culture est pris en compte par la variable *TCULT(t)* diminuée de la température minimale de développement, tc_{min}, caractéristique de **l'espèce**. La sénescence des feuilles est également simulée : elle dépend du temps et de **l'espèce et de la variété**, notamment au travers des paramètres *durvieI* (durée de vie des feuilles juvéniles) et *durvieF* (durée de vie des feuilles adultes).

- Le **module Croissance aérienne** calcule ensuite l'interception du rayonnement, *RAINT(t)*, selon une loi de Beer dont le coefficient d'extinction est une caractéristique de l'espèce. La vitesse de croissance potentielle en matière sèche, $\Delta MS(t)$, est une fonction du rayonnement intercepté, *RAINT(t)* :

$$\Delta MS(t)=[\varepsilon_{bmax}\times RAINT(t)-coef_b\times RAINT^2(t)]\times ftemp[TCULT(t)] \qquad \text{Équation II.3-2}$$

où *ftemp* est une fonction qui traduit l'effet de la température sur la conversion du rayonnement intercepté en biomasse et a une forme en cloche. ε_{bmax} est l'efficience maximale de conversion du rayonnement, caractéristique de l'espèce ; sa valeur dépend du stade de développement de la culture et correspond aux paramètres *efcroijuv*, *efcroiveg* et *efcroirepro* respectivement pour les phases juvénile, végétative et reproductive.

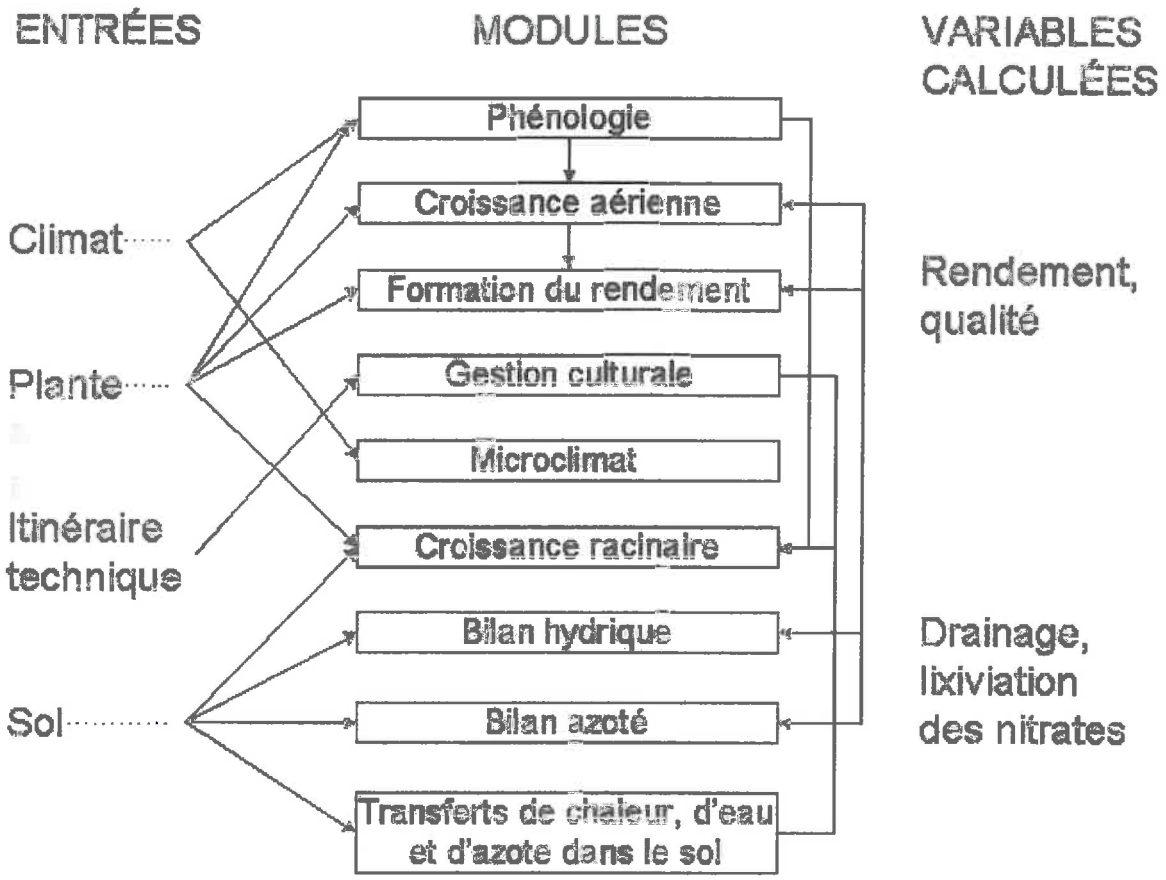


Figure II.3-2 : Entrées et sorties des neufs modules de STICS. Les flèches traduisent les principales influences respectives. Les flèches liant les modules et situées au centre de la figure symbolisent la croissance potentielle ; les flèches positionnées à droite des modules représentent l'effet de stress sur la croissance potentielle.

- Le **module Formation du rendement** calcule les composantes du rendement. La date de début du remplissage des grains est déterminée par un paramètre caractéristique de **l'espèce et de la variété**, *stlevdrp*. Le nombre de grains dépend linéairement du taux de croissance de la culture entre la fin de la croissance foliaire et le début de remplissage des grains selon les paramètres caractéristiques de l'**espèce** *cgrain* et *cgrainv0* et du nombre de grains maximal, *nbgrmax*, spécifique de la **variété**. La matière sèche des grains et la quantité d'azote accumulée dépendent d'indices de récolte croissant linéairement au cours du temps : *vitircarbT* et *vitirazo*, respectivement par rapport à la biomasse et à la quantité d'azote absorbée.

A ce stade de la description, tous les calculs effectués ne dépendent que du climat et des caractéristiques de la plante. Ceci correspond donc à une croissance potentielle de la culture (Figure II.3-2). La description de la croissance racinaire et la simulation des bilans hydrique et azoté sont nécessaires à la prise en compte de l'itinéraire technique et des interactions entre le sol et la plante qui se traduisent par des stress hydrique et azoté :

- Le **module Croissance racinaire** calcule la croissance du front racinaire, *ZRAC(t)*, qui dépend de **l'espèce et de la variété** par le biais du paramètre *croirac*, de la température et de l'humidité du sol ainsi que de sa densité apparente, soit des caractéristiques provenant tant du **climat** que du **sol**. L'influence de la densité apparente est traduite notamment par le paramètre *contrdamax*. En outre, la croissance s'arrête à un stade fourni par le module Phénologie et qui correspond à la fin de la production de *LAI*. Elle peut également être stoppée prématurément si le front racinaire rencontre un obstacle pédologique matérialisé par une variable caractéristique du **sol**, *OBSTARAC*. Une production de longueur racinaire est ensuite calculée selon un formalisme proche de celui utilisé pour *LAI* dans l'option choisie (« densité vraie »). Cette production est notamment affectée par la température du sol et l'anoxie. Les paramètres **plante** qui sont impliqués dans cette relation sont *draclong* et *lvfront*. La répartition de la longueur de racine produite est ensuite effectuée selon un formalisme qui tient compte de la répartition au jour précédent, de l'humidité du sol, de la densité apparente et de l'anoxie. Cette option de simulation a été choisie justement pour tenir compte de l'effet de la densité apparente sur l'enracinement. Enfin, la sénescence des racines intervient après une durée thermique de vie des racines, *stdebsenrac*. Ce module permet donc de connaître quotidiennement

sur des couches de 1 cm d'épaisseur les densités de racines, intégrées sur les cinq horizons *h* dans la variable *LRACH(h,t)*.

- Le **module Bilan hydrique** calcule, en fonction des caractéristiques de la **plante**, du **sol** et du **climat**, la teneur en eau dans chaque couche de sol de 1 cm. Il simule la transpiration de la plante, l'évaporation du sol, l'infiltration et la redistribution de l'eau dans le sol. Il permet ainsi de calculer les indices de stress hydriques.

- Le **module Bilan azoté** calcule l'offre du sol, bilan des fournitures du sol par minéralisation, des apports d'engrais et des pertes par organisation, émissions gazeuses, lixiviation des nitrates et absorption par la plante. Cette dernière est égale au minimum de l'offre du sol et de la demande de la culture. La demande est définie par le concept de courbe maximale de dilution d'azote. Pour calculer l'INN, qui déterminera les différents indices de stress azotés, le modèle utilise le concept de courbe critique de dilution de l'azote dans la plante.

- Le **module Transferts de chaleur, d'eau et d'azote dans le sol** est également impliqué dans ces calculs puisqu'il détermine les quantités d'eau et d'azote qui se trouvent à l'interface sol-racines.

- Trois indices de stress concernant l'alimentation en eau sont calculés : *TURFAC(t)* qui agit sur la croissance de *LAI*, *SWFAC(t)* qui affecte l'efficience de conversion et *SENFAC(t)* qui concerne la sénescence de *LAI*. Ces indices dépendent de seuils (Figure II.3-3-a), respectivement les variables *TETURG(t)*, *TETSTOMATE(t)* et *TETSEN(t)* (cette dernière étant proportionnelle à la première) dont le calcul implique les conditions hydriques du sol et dépend de deux paramètres spécifiques de l'**espèce** : *psisto* et *kmax*.

- Trois indices de stress azoté sont également calculés (Figure II.3-3-b) : *INNS(t)* qui diminue l'efficience de conversion, *INNLAI(t)* qui ralentit la croissance foliaire et enfin *INNSENES(t)* qui accélère la sénescence foliaire. Ces trois indices de stress sont calculés en fonction de l'INN respectivement grâce aux paramètres *innmin*, *innturgmin* et *innsen*. Le paramètre *innmin* marque la limite inférieure qu'ils peuvent atteindre.

- L'effet combiné des stress sur la croissance de *LAI* (Equation II.3-1) est simulé en multipliant son accroissement potentiel, *ΔLAI(t)*, par le minimum de *TURFAC(t)* et de *INNLAI(t)*, compris entre 0 et 1. L'indice de stress minimal entre *SENFAC(t)* et *INNSENEC(t)* diminue, également par multiplication, la durée de vie des feuilles. L'accroissement de matière sèche, *ΔMS(t)*, (Equation II.3-2), est diminué par le produit de *SWFAC(t)* et de *INNS(t)*.

Cette présentation de l'action des différents modules de STICS (mis à part le module Microclimat, non utilisé dans le cadre de notre étude) a permis de dégager les principaux paramètres θ et variables d'état du modèle qui seront impliqués dans son paramétrage.

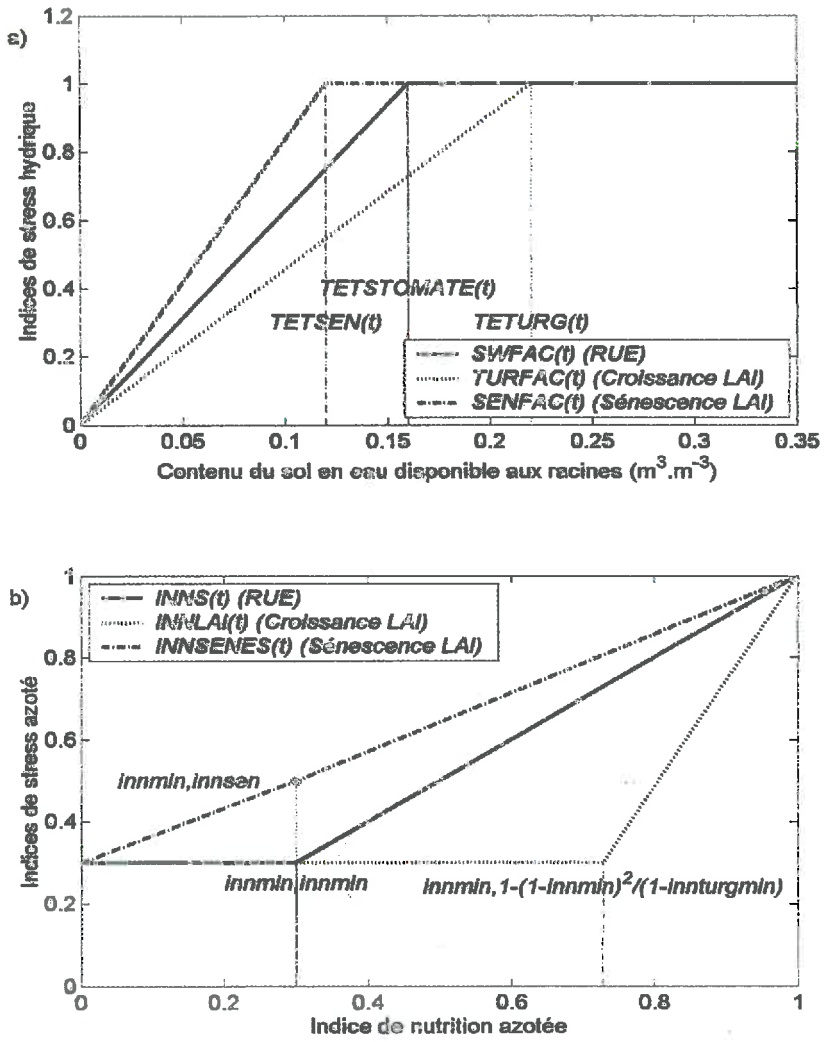


Figure II.3-3 : Calcul des indices de stress hydriques (a) et azotés (b) de STICS.

II.3.1.3 Relations entre variables d'entrées, d'état et de sortie.

Ce paragraphe présente de manière synthétique les relations entre les variables d'entrée, d'état et de sortie et se base sur le commentaire de la Figure II.3-4. Ce schéma servira également de support à l'explication de la méthodologie de paramétrage. Dans la suite du texte, les chiffres entre parenthèses renvoient aux chiffres entourés sur la figure.

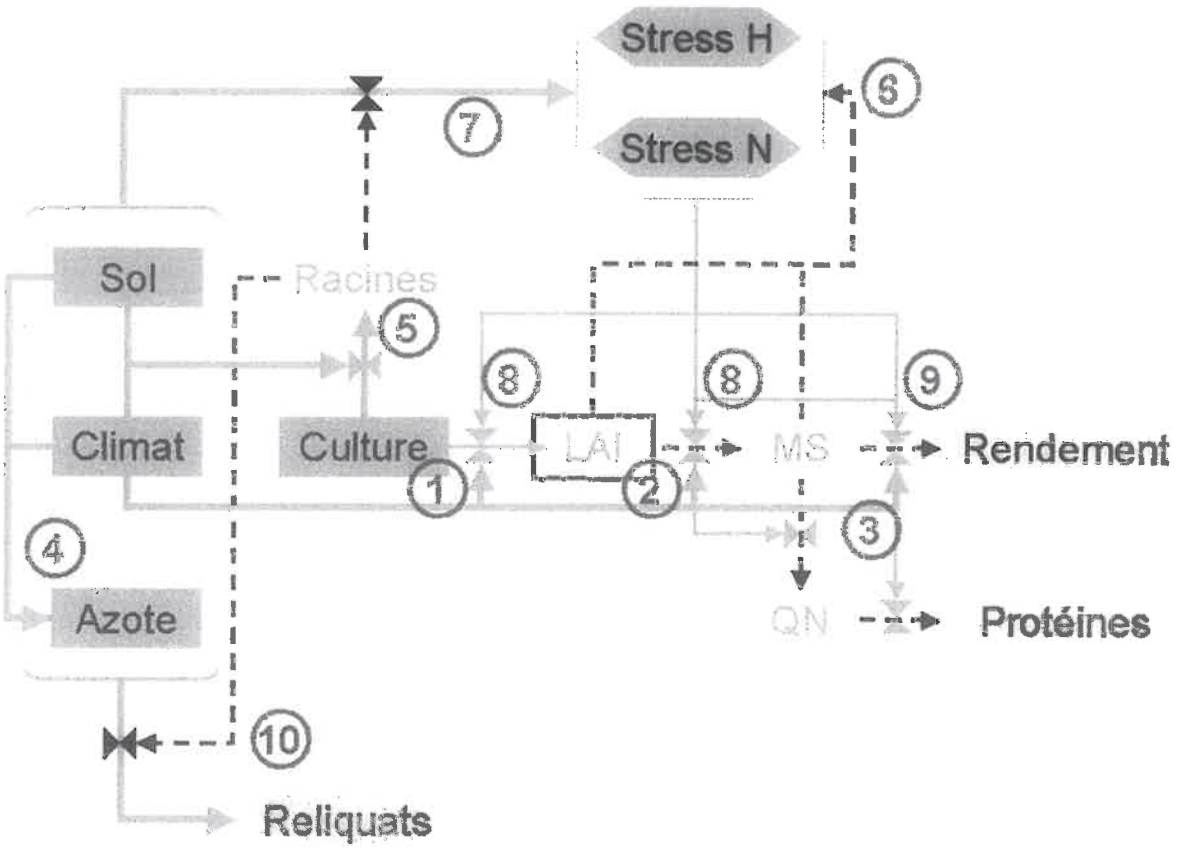


Figure II.3-4 : Relations entre les principales variables d'entrée, d'état et de sortie de STICS. Les numéros renvoient aux commentaires effectués dans le texte.

Les données **climatiques** permettent de calculer des stades de développement en fonction desquels la **culture** produit une surface foliaire potentielle (1). *LAI(t)* sert à générer de la matière sèche, *MS(t)* (2), qui à son tour détermine le rendement en grains, *G(t)*. La teneur en protéines, *P(t)*, des grains (teneur en azote × 5.7 ; Grundy *et al.*, 1996) dépend de la courbe d'absorption potentielle d'azote (courbe de dilution maximale) (3). Le calcul des stress hydrique et azoté, qui affectent ces variables d'état, et des autres variables implique l'utilisation d'autres entrées : le **sol** et l'itinéraire technique, principalement les **apports d'azote minéral**. Les caractéristiques du sol et du climat déterminent la quantité d'azote minéral présente dans le sol (4). Parallèlement, la culture émet des racines en fonction de la

phénologie et des propriétés du **sol** (5). Les variables *LAI(t)* et *MS(t)* conditionnent la demande en eau et en azote de la culture (6). La demande en eau du système sol-culture est également liée aux propriétés du sol. La quantité d'eau et d'azote que peut absorber la plante dépend de leur disponibilité dans le sol et de la densité de racines (7). Les indices de stress hydrique et azoté ainsi calculés affectent le développement potentiel de *LAI(t)* et de *MS(t)* (8) ainsi que des variables de sorties, *G(t)* et *P(t)* (9). La dernière variable de sortie qui nous intéresse, le reliquat d'azote minéral à la récolte, *R(t)*, résulte de la différence entre les apports dans le sol au cours de la culture (quantité d'azote initial, minéralisation et apports extérieurs) et les prélèvements par les racines (10).

II.3.1.4 Paramètres importants

Les paramètres importants cités dans les paragraphes ci-dessus sont ceux qui ont été retenus par l'analyse de sensibilité de STICS par Ruget *et al.* (2002) et, eu égard au changement de version, par expertise (Justes, comm. pers.). Le Tableau II.3-1 synthétise ces paramètres et les associe au module dans lequel ils sont impliqués et à la variable principale qu'ils contrôlent.

Paramètre	Module	Variable
stlevamf, stamflax, spltger	Phénologie	*LAI(t)*
$dlai_{max}$, $pentlai_{max}$, $vlai_{max}$, tc_{min}	Croissance aérienne	*LAI(t) (croissance)*
durvieI, durvieF	Croissance aérienne	*LAI(t) (sénescence)*
efcroijuv, efcroiveg, efcroirepro	Croissance aérienne	*MS(t)*
stlevdrp, cgrain, cgrainv0, nbgrmax, vitircarbT	Formation du rendement	*G(t)*
vitirazo	Formation du rendement	*P(t)*
croirac, contrdamax	Croissance racinaire	*ZRAC(t)* (front racinaire)
draclong, lvfront, stdebsenrac	Croissance racinaire	*LRACH(t,h)* (densité de racines)
psisto, kmax	Bilan hydrique	*Stress hydriques*
innmin, innturgmin, innsen. innmin	Bilan azoté	*Stress azotés*

Tableau II.3-1 : Paramètres les plus influents de STICS et module dans lequel ils sont impliqués.

L'utilisation de ces paramètres sera expliquée au cours de la partie II.3.2.

II.3.1.5 Les variables d'entrée sol et climat

II.3.1.5.1 Variables climatiques

Le Tableau II.3-2 donne les variables climatiques qui sont nécessaires au fonctionnement du modèle. Elles doivent être acquises à un pas de temps quotidien.

Variables	Signification
$T_{MIN}(t)$	Température minimale (°C)
$T_{MAX}(t)$	Température maximale (°C)
$R_G(t)$	Rayonnement global journalier (MJ m^{-2} j^{-1})
$ETP(t)$	Evapotranspiration potentielle (mm j^{-1})
$Pluies(t)$	Précipitations (mm j^{-1})

Tableau II.3-2 : Variables climatiques utilisées dans STICS.

II.3.1.5.2 Variables sol

A l'échelle d'un cycle cultural, les caractéristiques pédologiques du sol ne dépendent pas du temps. Mais comme elles varient spatialement, conformément à la définition proposée dans le Chapitre I, elles seront considérées à ce titre comme des variables. Le Tableau II.3-3 les décrit et indique le type de stress qu'elles conditionnent le cas échant. Celles qui sont affublées d'un *(h)* désignent les variables à renseigner pour chaque horizon ; les autres ne concernent que le premier horizon ou l'ensemble du profil (comme *OBSTARAC*).

Variables	Signification	Stress
ARGI	Teneur en argile de l'horizon de surface (%)	H
NORG	Teneur en azote organique de l'horizon de surface (%)	N
PROFHUM	Profondeur d'humification (cm)	N
CALC	Teneur en calcaire de l'horizon de surface (%)	H et N
pH	pH de l'horizon de surface	
ALBEDO	Albédo de l'horizon de surface sec	
q0	Seuil de cumul d'évaporation jusqu'auquel l'évaporation est potentielle (mm)	H
OBSTARAC	Profondeur du sol à laquelle un obstacle physique ou chimique à l'enracinement apparaît (cm)	H et N
EPC(h)	Epaisseur de l'horizon *h* (jusqu'à 5 horizons)	
HCCF(h)	Teneur en eau à la capacité au champ de l'horizon *h* (%)	H
HMINF(h)	Teneur en eau au point de flétrissement permanent de l'horizon *h* (%)	H
DA(h)	Densité apparente de l'horizon *h* (g cm^{-3})	H et N
CAILLOUX(h)	Teneur volumique en cailloux de l'horizon *h* (%)	
DA_CAILLOUX(h)	Masse volumique apparente des cailloux (g cm^{-3})	
HCC_CAILLOUX(h)	Teneur en eau à la capacité au champ des cailloux (%)	

Tableau II.3-3 : Variables sol utilisées dans STICS et les stress qu'elles conditionnent : H pour le stress hydrique et N pour le stress azoté.

II.3.1.6 Quelques résultats obtenus auparavant avec STICS

Le Tableau II.3-4 présente des résultats d'évaluation du modèle obtenus antérieurement avec la version 4 sur vingt-deux situations très variées (Brisson *et al.*, 2002).

La comparaison des RMSEs et RMSEu montre que, sur l'ensemble du jeu de données très varié utilisé pour apprécier ces valeurs, le modèle est peu biaisé et qu'en revanche la plus grande part de l'erreur commise provient de la dispersion et non du biais. Ceci correspond à la vocation de robustesse du modèle. Ces résultats seront à comparer avec ceux obtenus dans la suite de ce travail.

Statistique	*G* ($t\ ha^{-1}$)	*MS* ($t\ ha^{-1}$)	*P* (%)	*R* ($kg\ ha^{-1}$)	*B* ($kg\ ha^{-1}$)
RMSE	1.35	2.66	2.05	44	18.9
RRMSE	18 %	17 %	21 %	49 %	28 %
RMSEs	0.28	1.62	0.28	29	4.8
RMSEu	1.32	2.11	1.32	34	18.2

Tableau II.3-4 : Critères statistiques d'évaluation de STICS sur blé tendre pour 22 situations ; *G* : rendement en grains à 0% d'humidité ; *MS* : matière sèche à la récolte ; *P* : teneur en protéines des grains ; *R* : reliquats d'azote à la récolte ; *B* : bilan d'azote (*G*, *MS*, *R* d'après Brisson *et al.*, 2002 ; *P* et *B* d'après Brisson, comm. pers., tirée du jeu de données ayant servi à écrire l'article cité).

II.3.2 Méthodologie de paramétrage de STICS

Cette partie présente la méthodologie utilisée pour réaliser le paramétrage de STICS : elle s'appuie sur l'analyse du modèle qui vient d'être présentée et dégage les caractéristiques essentielles auxquelles devra répondre la base de données utilisée pour ce faire. Comme il s'agit d'estimer les paramètres du fichier plante, les variables d'entrées concernant le sol et le climat doivent être fiables.

Chaque modèle de culture possédant ses spécificités, il n'existe pas de méthodologie générale de paramétrage (Launay, comm. pers.). Nous avons donc adapté la méthode initiée par Ghiloufi (1999) et développée par Dorsainvil (2002) pour la version 5 à notre situation (Justes, comm. pers.). Contrairement à ces travaux, il ne s'agissait pas ici d'adapter STICS à une nouvelle culture, mais une phase de paramétrage était nécessaire pour deux raisons : (i) le changement de version et de formalismes importants comme celui assurant la simulation de *LAI* ; (ii) l'adaptation à la variété *Shango*, utilisée dans l'expérimentation « agriculture de précision » de Chambry.

II.3.2.1 Présentation générale de la méthode

Il existe deux grandes façons d'aborder l'estimation des paramètres d'entrée d'un modèle : (i) une analyse préalable du modèle et une estimation séquentielle des paramètres par module ; (ii) une approche statistique globale. La méthode utilisée ici fait partie de ce premier ensemble.

Malgré les lacunes que représente d'un point de vue statistique une telle méthode, elle a été utilisée pour plusieurs raisons : (i) elle avait déjà été employée avec STICS ; (ii) vu le temps imparti à cette étape du travail, il n'était pas envisageable de développer un outil

informatique adapté à STICS qui aurait permis de mettre en œuvre des méthodes par approche statistique globale telles que celle proposée par Wallach *et al.* (2001) ; (iii) l'approche de paramétrage module par module basée sur l'analyse du fonctionnement du modèle et sur la connaissance des phénomènes demeure intéressante et informative, malgré le fait que cela implique une estimation séquentielle des paramètres.

L'outil utilisé pour réaliser le paramétrage est un programme fourni avec STICS et codé en FORTRAN®, basé sur la méthode du simplexe (*cf.* Chapitre III pour la description de l'algorithme du simplexe). Il permet d'optimiser jusqu'à cinq paramètres sur une ou plusieurs variables d'état. Le critère optimisé est un critère des moindres carrés pondérés. La pondération est cependant unique pour une même variable, quelle que soit la date de mesure de la variable. Il s'écrit de la façon suivante :

$$C(\theta)=\frac{1}{N.D}\sum_{n=1}^{N}\sum_{d=1}^{D}\frac{(Y_{n,d}-f(X;\theta;t_d))^2}{p_n} \qquad \text{Équation II.3-3}$$

où $C(\theta)$ désigne la valeur du critère à minimiser pour le jeu de valeurs de paramètres θ testé, N le nombre de variables d'état considérées, D le nombre de dates de mesures, $Y_{n,d}$ la valeur observée à la date d de la variable n, $f(X;\theta;t_d)$ la valeur simulée à la date d de la variable n pour le jeu de valeurs de paramètres θ et p_n le facteur de pondération spécifique de la variable n.

La méthode proposée comporte trois étapes :

- la première utilise la version de STICS qui permet de forcer *LAI* à partir de mesures : dans cette version, cette variable n'est pas calculée mais devient une entrée journalière, ce qui implique d'interpoler *LAI* entre les dates de mesures disponibles à l'aide d'un modèle simple de cinétique (Baret, 1986). Ceci permet de réaliser l'optimisation des paramètres des modules dépendant de cette variable capitale sans faire d'erreur sur sa prévision. A ce stade, la date de levée est également forcée par la valeur mesurée ;

- la deuxième utilise la version normale de STICS, qui simule *LAI*. Cette étape permet d'optimiser les paramètres qui régissent la mise en place de *LAI* ;

- la dernière étape est une optimisation inter-modules sur les cinq paramètres jugés les plus sensibles d'après analyse des étapes précédentes.

II.3.2.2 Première étape : *LAI* forcé

La Figure II.3-5 illustre cette étape du paramétrage.

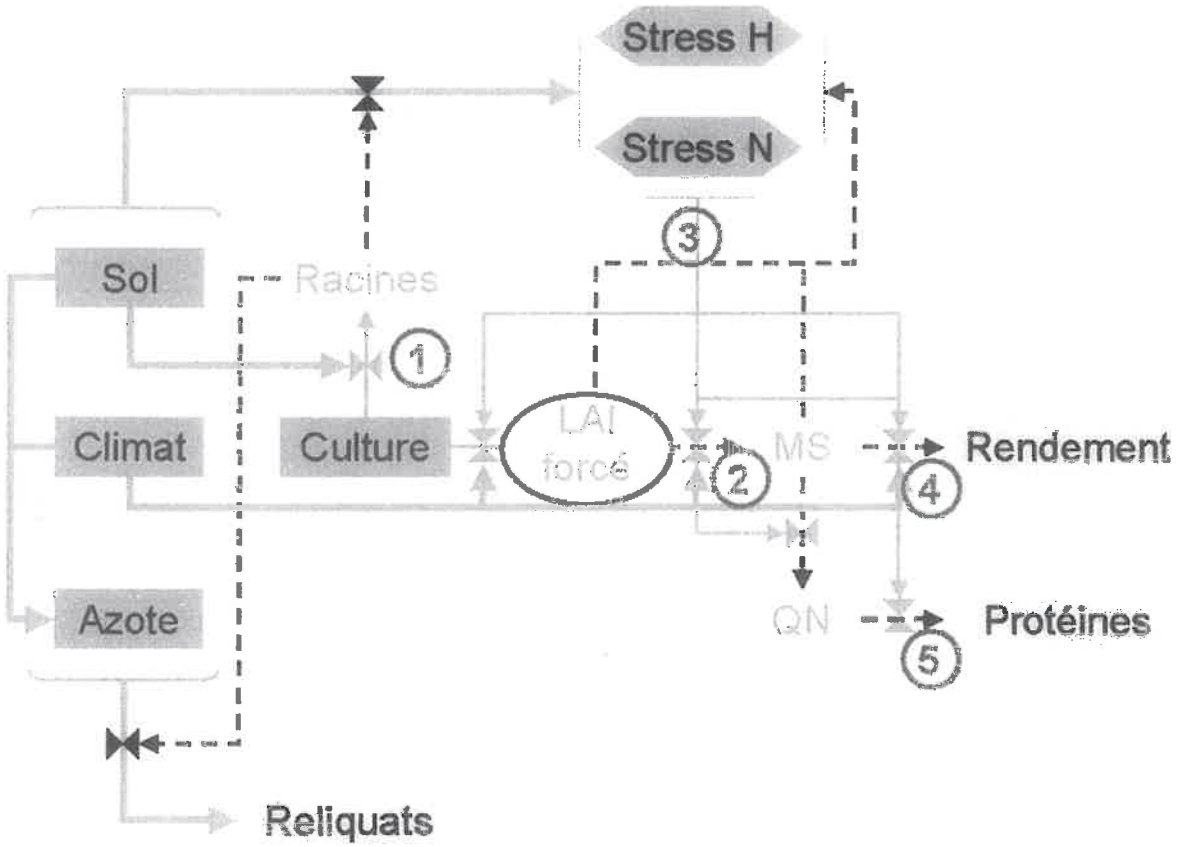


Figure II.3-5 : Phases de paramétrage de STICS avec *LAI* et la date de levée forcée. Les numéros renvoient aux commentaires effectués dans le texte.

Il est important de noter que les jeux de données utilisés à cette étape doivent comporter au moins cinq observations de *LAI* afin de pourvoir ajuster le modèle simple de *LAI* qui comporte cinq paramètres.

II.3.2.2.1 Estimation des paramètres de croissance racinaire

Il est essentiel de paramétrer en premier ce module car c'est lui qui conditionne l'alimentation azotée et hydrique de la culture et par là la croissance (Figure II.3-5) (1). Les paramètres estimés à cette étape sont les suivants : *draclong*, *contrdamax* et *lvfront*. La variable cible est *LRACH(h,t)*. On estime également le paramètre *croirac* par rapport à la variable *ZRAC(t)*.

Le jeu de données nécessaire à cette phase doit comporter des observations sur les racines, sans stress.

II.3.2.2.2 Estimation des paramètres d'efficience de conversion

On estime d'abord le paramètre *efcroijuv* qui correspond à la phase juvénile. La variable cible est *MS(t)* au cours de cette phase. Pour réaliser cette phase, il est donc nécessaire de posséder un jeu de données comportant de nombreuses observations de matière sèche précoce. On estime ensuite *efcroiveg* et *efcroirepro* qui correspondent respectivement aux efficiences de conversion lors des phases végétative et reproductive. Des jeux de données possédant de nombreuses observations de matière sèche lors de ces phases sont donc nécessaires (Figure II.3-5) (2). Ces jeux de données doivent être en conditions potentielles, donc sans stress.

II.3.2.2.3 Estimation des paramètres de stress hydrique

Une fois les paramètres gouvernant la croissance potentielle fixés, il convient d'estimer les paramètres qui interviennent dans le déterminisme des variables de stress hydrique (Figure II.3-5) (3). Les paramètres *psisto* et *kmax* sont ajustés sur la variable *MS(t)* grâce à des jeux de données comportant des épisodes de carences en eau. Les paramètres de stress azotés ne seront pas optimisés à cette étape : nous avons choisi d'agir sur ceux qui concernent *LAI* et qui seront donc abordés à la deuxième étape. Quant aux paramètres d'absorption de l'azote par la plante, nous avons considéré qu'ils étaient connus pour le blé, étant donné que le formalisme n'a pas été modifié entre les deux versions.

II.3.2.2.4 Estimation des paramètres de formation du rendement

Les paramètres impliqués dans le formalisme décrivant la formation du rendement, *vitircarbT*, *cgrain* et *cgrainv0*, ont été estimés en ajustant la variable *G(t)*. Les paramètres *stlevdrp* et *nbgrmax* ont été fixés d'après des caractéristiques variétales, pour *Shango*, 837 °C j^{-1} et 30 000 respectivement. Ici encore, comme nous travaillons toujours sous l'hypothèse des conditions potentielles, les jeux de données utilisés doivent correspondre à des situations sans stress (Figure II.3-5) (4). Le paramètre *vitirazo* est ensuite optimisé par rapport à la variable *P(t)* dans des conditions similaires (Figure II.3-5) (5).

II.3.2.3 Deuxième étape : *LAI* simulé

La description de cette partie s'appuie sur la figure suivante :

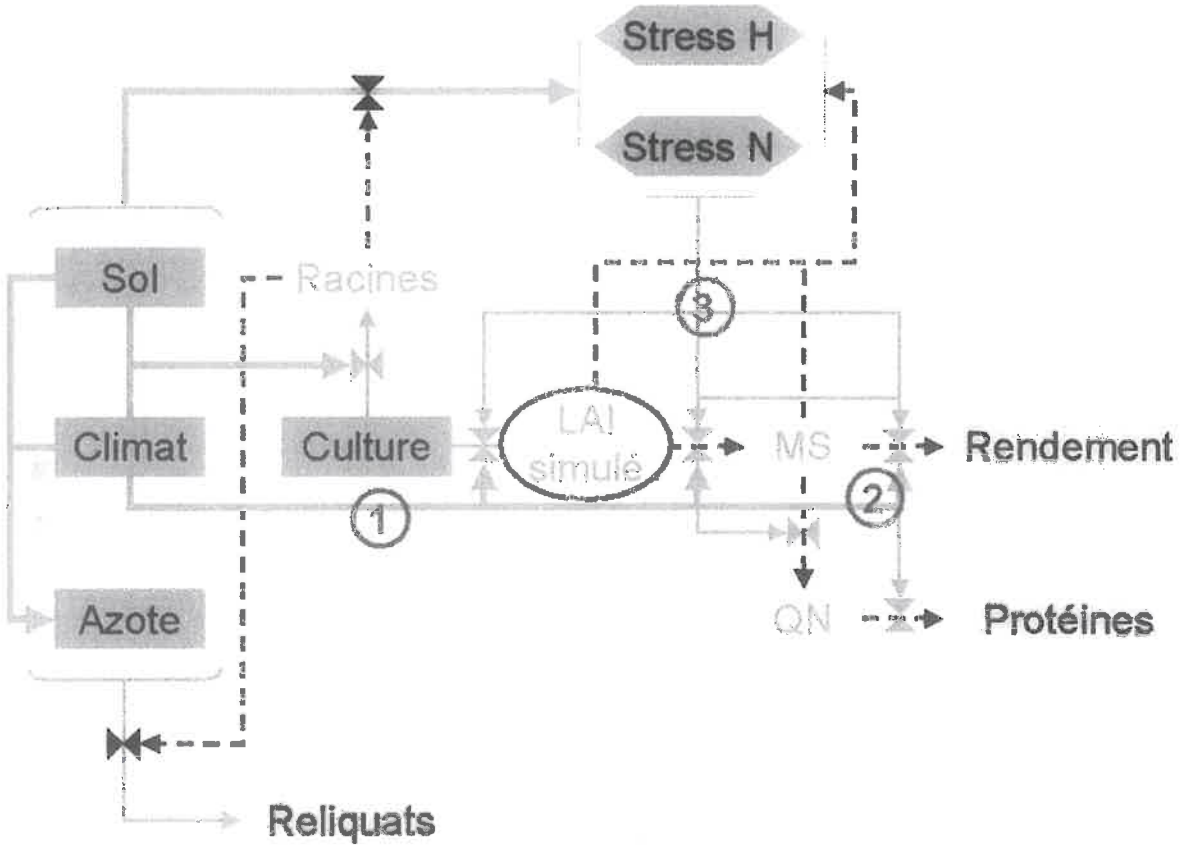


Figure II.3-6 : Phases de paramétrage de STICS avec *LAI* et la date de levée simulés. Les numéros renvoient aux commentaires effectués dans le texte.

II.3.2.3.1 Estimation de la date de levée

Le paramètre *stpltger* est estimé grâce à des observations de dates de levée. La culture à l'origine du jeu de données utilisé ne doit pas avoir subi de conditions exceptionnelles à la levée (Figure II.3-6) (1).

II.3.2.3.2 Estimation de la durée des stades de développement

Les paramètres *stlevamf* et *stamflax*, qui sont caractéristiques de la variété, doivent être optimisés par rapport à des observations de stades de développement, en assimilant le stade AMF de STICS au stade épi à 1 cm et le stade LAX au stade gonflement, sortie dernière feuille (Brisson *et al.*, 2002). La culture ne doit pas avoir subi de conditions induisant un retard de développement. Les mesures des caractéristiques connues de variétés proches ont été utilisées comme première approximation des valeurs de ces paramètres. De la même façon, la

durée entre la levée et le début de remplissage des grains, *stlevdrp*, est estimée à partir d'observation de dates de début de remplissage des grains (floraison), Figure II.3-6 (2).

II.3.2.3.3 Estimation des paramètres de la cinétique de *LAI*

Les paramètres $vlai_{max}$, $pentlai_{max}$, et $dlai_{max}$ sont estimés en s'appuyant sur des données observées de *LAI(t)*. Les données doivent provenir d'expérimentations n'ayant pas subi de stress hydrique ou azoté. Il est préférable de n'utiliser que des essais réalisés avec la variété étudiée (*Shango*), même si ces paramètres sont peu dépendants de la variété. Les jeux de données doivent comporter de nombreuses observations de *LAI*, notamment des observations précoces. L'idéal serait de pouvoir estimer ces trois paramètres sur des données de *LAI* observé généré, c'est-à-dire donnant la surface foliaire totale produite, incluant la sénescence. La sénescence simulée par STICS serait alors neutralisée.

La sénescence doit ensuite être paramétrée : on revient alors à des observations de *LAI* net, incluant ses effets. Les paramètres concernés sont : *durvieI* et *durvieF*. En l'absence de mesures distinctes de *LAI* produit et sénescent, c'est la variable *LAI* net qui sera utilisée, ce qui conduit à estimer en même temps les paramètres de l'élaboration du *LAI* et de sa sénescence.

II.3.2.3.4 Estimation des paramètres du stress azoté sur le *LAI*

Les paramètres *innturgmin et innsen* peuvent être estimés par rapport à *LAI* en utilisant des essais avec des carences azotées. Il est préférable d'intégrer une assez large gamme d'intensité de carence (Figure II.3-6) (3).

II.3.2.4 Troisième étape : estimation inter-modules

L'estimation des paramètres module par module peut conduire à des erreurs dues à la non prise en compte de certaines des interactions entre modules. Il est donc nécessaire de procéder à une seconde phase d'estimation de certains paramètres jugés sensibles et particulièrement importants en utilisant l'ensemble des jeux de données disponibles pour le paramétrage.

Il ressort de la présentation de cette méthode la nécessité de structurer la base des données utilisables en fonction des critères que nous avons énoncés : variables intermédiaires mesurées (*LAI* et *MS* notamment) et conditions de culture. Les essais qui ne répondent pas à

ces critères peuvent être conservés pour la validation. La partie suivante décrit les jeux de données sur lesquels nous nous sommes appuyés.

II.3.3 Données expérimentales

Les données expérimentales présentées ci-dessous serviront à la fois à :

- paramétrer et valider le modèle ;
- comparer des critères agro-environnementaux ;
- évaluer l'aptitude du modèle à recommander des doses d'azote et les conséquences de ces recommandations.

II.3.3.1 Présentation générale des essais utilisés

Toutes les expérimentations utilisées sont des essais de dose d'azote réalisés par Arvalis-Institut du végétal ou l'INRA. Les sites sont : Boigneville, (Essone), Thibie (Marne), La Minière (Yvelines), La Cheppe (Marne), Chambry (Aisne), Grignon (Yvelines). Ces essais peuvent être considérés représentatifs des conditions de production agricole au nord du Bassin Parisien.

Lieu	Années	Nombre d'essais	Nombre de traitements	Type de sol
Boigneville (91)	1998-1999	3 × 2	5	Limon à cailloux
Thibie (51)	1994-1995-1998-1999	1 × 4	7 à 10	Calcaire
La Minière (78)	2002	1	9	Limon profond
La Cheppe (51)	2002	1	6	Calcaire
Chambry (02)	2000-2001	1 × 2	6	Sablo-limoneux
Grignon (78)	1992-1995	1 × 2	1 utilisé / an	Limon

Tableau II.3-5 : Caractéristiques principales des essais de doses d'azote utilisés.

Le Tableau II.3-5 indique pour chaque site les années de récolte concernées, le nombre d'essais effectués, le nombre de traitements réalisés par essai et le type de sol. L'expérimentation de Boigneville comporte en réalité 15 essais azote, deux années de suite,

sur une même parcelle très hétérogène du point de vue de la pédologie. Trois essais seulement ont été utilisés par année afin de ne pas donner une trop grande prépondérance à ce site.

II.3.3.2 Mesure des variables d'entrée du modèle

Le Tableau II.3-6 renseigne les principales caractéristiques des variables d'entrée du modèle par site. Chacun d'eux comporte au minimum 5 doses d'azote différentes, dont certaines peuvent être identiques mais apportées selon un fractionnement différent.

Afin de permettre d'évaluer les caractéristiques climatiques lors de la phase de croissance et de maturation, les cumuls de températures moyennes et de précipitations sont donnés par année entre le 1er mars et le 30 juin. A titre de comparaison, les mêmes indicateurs mesurés pour l'année 1976, réputée pour avoir été particulièrement sèche (Jamagne, 1977), à la station de Météo France de Roupy (02) valent respectivement 1330 °C et 68 mm. Les moyennes pour la même station des 32 dernières années sont 1298±86 °C et 236±59 mm ; les moyennes des 10 dernières années sont 1365±78 °C et 224±58 mm. Seul l'essai de Boigneville 98 correspond à des conditions assez sèches. Pour la plupart des années, les cumuls de température sont plutôt supérieurs à la moyenne. L'accent a été mis sur la variété *Shango* dans le jeu de données afin de bien paramétrer cette variété, qui est utilisée sur le site expérimental de Chambry (*cf. infra*). Mais le changement de version a imposé une adaptation plus complète des paramètres plante : d'autres variétés ont été introduites pour traiter les paramètres non variétaux.

Il serait fastidieux de conter par le menu la façon dont les différentes variables d'entrée sol ont été renseignées pour chaque essai. Dans la majorité des cas, les données d'entrée 'sol' étaient lacunaires ; on peut notamment citer l'absence de mesures de teneurs en eau et en azote au semis pour les essais de La Cheppe et de La Minière. Dans le cas de Boigneville où tous les essais étaient situés sur une seule parcelle hétérogène (Laurent 2000), un important dispositif de mesure avait été mis en place, mais les données mesurées ne correspondaient pas forcément aux entrées de STICS. La variabilité de la parcelle et le fait que les mesures n'avaient pas été réalisées sur l'emplacement des essais rendaient l'extrapolation difficile. Pour Thibie et La Cheppe, les valeurs de densités apparentes ont été extrapolées de celles mesurées sur le site proche du Rafidin. Les données de teneurs en eau à la capacité au champ et au point de flétrissement permanent ont été déduites de séries de mesures de teneur en eau effectuées respectivement en milieu d'hiver et à la récolte. Toutes

Site-année	N° essais	Doses N (kg ha^{-1})	Cumuls de mars à juin		Variété	Données sur le sol
			T (°C)	P (mm)		
Boigneville 98	1 à 3	0-120-160-**200**-240	1511	150	*Soissons*	Parcelle
Boigneville 99	4 à 6	0-120-160-**200**-280	1557	239	*Soissons*	Parcelle
Thibie 94	7	80-120-160-200-**240**-280-320	1400	210	*Forby*	Région
Thibie 95	8	0-80-120-160-200-**240**-280-320	1266	244	*Apollo*	Région
Thibie 98	9	0-100-135-150-**200**-210-250-300-350-400	1508	239	***Shango***	Région
Thibie 99	10	0-100-135-150-200-**210**-250-300-350-400	1535	215	***Shango***	Région
La Minière 02	11	0-110-150-**190**-230-*190-230-190-190*	1545	187	***Shango***	Mixte
La Cheppe 02	12	0-140-**220**-300-*220-180*	1507	195	***Shango***	Région
Chambry 00	13	0-60-120-180-**240**-300	1507	309	***Shango***	Parcelle
Chambry 01	14	0-120-160-**220**-240-280	1407	320	***Shango***	Parcelle
Stations Ch 00		240 (4 stations)	1507	309	***Shango***	Parcelle
Grignon 92		260	1476	211	*Soissons*	Inconnue
Grignon 95		300	1374	203	*Soissons*	Inconnue

Tableau II.3-6 : Principales informations concernant les variables d'entrée du modèle par site-année. « Stations Ch 00 » désigne des stations de calibration suivies à Chambry en 2000 avec un suivi racinaire. La colonne « numéros des essais » donne pour chaque site-année les numéros attribués aux essais de doses d'azote. Les doses indiquées en gras correspondent à la dose bilan préconisée par AZOBIL, en italique à des apports réalisés avec un fractionnement différent. La colonne « Données sur le sol » traduit l'échelle à laquelle l'information sur les données sol a été acquise : 'Parcelle' signifie que les mesures ont été réalisées de façon locale (pas forcément sur l'essai proprement dit cependant), 'Région' signifie que les informations ont dû être extraites de bases de données de la même région, 'Mixte' signifie que les mesures spécifiques au premier horizon (*ARGI, NORG, PROFHUM, CALC*) ont été faites sur la parcelle, les autres ont été extrapolées à partir de données régionales. L'information n'était pas communiquée avec le jeu de données de Grignon.

ces mesures ne correspondaient pas aux parcelles sur lesquelles avaient été effectués les essais. A La Minière, les données principales concernant le premier horizon ont été mesurées sur la parcelle même, mais les humidités caractéristiques des différents horizons sont celles d'un sol limoneux profond typique. A Chambry, les mesures essentielles ont été faites sur la parcelle (*cf. infra*), mais pas aux emplacements mêmes des essais azote.

Il ressort de ce bref descriptif que les valeurs des paramètres sol, pourtant essentielles au paramétrage et à la validation du modèle, étaient souvent incertaines. Au cours de la phase même de validation, certaines ont été remises en question et modifiées.

II.3.3.3 Mesure des variables d'état et partition du jeu de données

Le Tableau II.3-7 indique les variables qui ont été mesurées par essai et, le cas échéant, le nombre de dates de mesure (pour *LAI* et la matière sèche, notée *MS*).

Essai	*LAI*	*MS*	Racines	Reliquats	Rendement	Protéines	Stades
Boigneville 98	-	-	-	Oui	Oui	Oui	-
Boigneville 99	-	-	-	Oui	Oui	Oui	-
Thibie 94	-	-	-	-	Oui	Oui	-
Thibie 95	-	-	-	Oui	Oui	Oui	-
Thibie 98	-	-	-	-	Oui	Oui	-
Thibie 99	-	-	-	-	Oui	Oui	-
La Minière 02	3*	4	-	-	Oui	Oui	-
La Cheppe 02	3*	4	-	Oui	Oui	Oui	-
Chambry 00	7	8	-	-	Oui	Oui	Oui
Chambry 01	8	9	-	-	Oui	Oui	Oui
Stations Ch 00	8	9	Profils	Oui	Oui	Oui	Oui
Grignon 92	16	17	-	-	Oui	Oui	Oui
Grignon 95	22	23	-	-	Oui	Oui	Oui

Tableau II.3-7 : Variables mesurées ou observées par essai. Les mesures de *LAI* marquées d'une astérisque ont été réalisées au LICOR LAI-2000®. Les tirets indiquent l'absence de mesure.

Les mesures de *LAI* ont été effectuées sur des prélèvements de végétaux par mesure de la surface des feuilles au laboratoire grâce à un planimètre, sauf à La Cheppe et à La Minière où elles ont été réalisées sur le terrain à l'aide d'un LICOR LAI-2000® par mesure indirecte. Les mesures de matière sèche ont toujours été établies par prélèvement et pesée. Les profils de racines sur les stations de Chambry ont été dressés par la méthode de Tardieu & Manichon (1986). Une grille de 1m×1m est plaquée sur la paroi d'une fosse. On note la présence ou l'absence de racines dans chaque maille de 4 cm^2. Un calcul simple permet de convertir la fréquence des impacts de racines en une densité, variable simulée par STICS : *LRACH(h,t)*. On note aussi la profondeur maximale d'enracinement, équivalent de *ZRAC(t)*.

Seuls les essais de Chambry et de Grignon comportent les dates de levées et un nombre suffisant d'observations de *LAI* et de matière sèche pour être utilisables pour le paramétrage. Les autres essais seront utilisés pour la validation. On note que le nombre de sites permettant la validation des simulations du reliquat d'azote à la récolte est moins important que pour le rendement et la teneur en protéines.

Le Tableau II.3-8 reprend les étapes de la méthode de paramétrage en indiquant quelles parties du jeu de données ont été utilisées pour chacune d'elles.

Etapes	Phases	Conditions requises	Essais utilisés
LAI forcé	Croissance racinaire	0 stress, obs. racines	St 00
	efcroijuv	0 stress, obs. précoces	St 00, Gr 95
	efcroiveg, efcroirepro	0 stress	Ch 00+, St 00, Gr 92&95
	psisto, kmax	Stress H, 0 stress N	*Aucune satisfaisante*
	Formation du rendement	0 stress	Ch 00+, St Ch 00, Ch 01+
LAI simulé	*stpltger*	0 stress, obs. levée	Ch 00-5, Ch 01-5
	stlevamf, stamflax	0 stress, obs. stades	Ch 00-5, Ch 01-5
	stlevdrp	0 stress, obs. stades	Ch 00-5, Ch 01-5
	Cinétique de *LAI*	0 stress, *Shango*	Ch 00+, St Ch 00, Ch 01+
	Stress azoté sur *LAI*	Stress N, 0 stress H	Ch 00, Ch 01

Tableau II.3-8 : Etapes du paramétrage de STICS, conditions requises sur les données et jeux de données utilisés correspondant. Obs. : observations ; St 00 : les stations de Chambry 2000 ; Gr : Grignon ; Ch : Chambry ; Ch 00+ : les 3 plus fortes doses de l'essai azote de Chambry 2000 ; Ch 00-5 : la 5ème dose de l'essai azote de Chambry 2000.

L'ensemble de ces données à l'exception de celles de Grignon et des stations de calibration de Chambry (qui ne sont pas des essais doses d'azote) serviront à comparer des critères agro-environnementaux et à évaluer l'aptitude du modèle à réaliser des préconisations de fertilisation azotée.

II.4 Paramétrage de STICS et évaluation de sa réponse à la fertilisation azotée

Cette partie est organisée en trois paragraphes : le premier donne les résultats du paramétrage, le deuxième sa validation, ce qui permet de caractériser la réponse de STICS à la fertilisation azotée. Le dernier synthétisera les points importants qui sont à retenir de cette partie.

II.4.1 Résultats du paramétrage

II.4.1.1 Valeurs des paramètres obtenues

Le Tableau II.4-1 donne les résultats obtenus pour les paramètres ajustés et indique à titre de comparaison les valeurs de la version 4 standard de STICS.

Paramètre	STICS 4	STICS 5	Paramètre	STICS 4	STICS 5
draclong	80 *	29	*nbgrmax*	variétal	22 500
contrdamax	nouveau	0.3	*vitirazo*	$1.3\ 10^{-2}$	$1.76\ 10^{-2}$
lvfront	$5\ 10^{-4}$ *	0.0814	*stpltger*	50	110
efcroijuv	2.2	2.87	*stlevamf*	variétal	235
efcroiveg	4.25	4.14	*stamflax*	variétal	340
efcroirepro	4.25	3.80	*stlevdrp*	variétal	837
psisto	15	15	$vlai_{max}$	nouveau	2
kmax	1.2	1.0	$pentlai_{max}$	nouveau	5.5
vitircarbT	$7.0\ 10^{-4}$	$7.7\ 10^{-4}$	$dlai_{max}$	$4.4\ 10^{-4}$	$4.4\ 10^{-4}$
cgrain	1043 *	0.064	*durvieF*	nouveau	183
cgrainv0	-4160 *	-0.1429	*innturgmin*	nouveau	-0.2
			innsen	nouveau	0.35

Tableau II.4-1 : Comparaison des valeurs des paramètres ajustés de la version 4 standard de STICS et celles obtenues pour STICS 5 par paramétrage. 'variétal' : le paramètre concerné a été ajusté ou mesuré pour la variété *Shango* et l'équivalent n'existait pas dans STICS 4 ; 'nouveau' : n'existait pas en tant que paramètre dans STICS 4 ; * : l'unité et le formalisme de ces paramètres ont été modifiés.

II.4.1.2 Simulation des variables et problèmes rencontrés

Les résultats du paramétrage vont être montrés à travers les principales variables ajustées, *LAI(t)*, *MS(t)* et *G(t)*.

La Figure II.4-1 et la Figure II.4-2 illustrent les problèmes rencontrés au cours du paramétrage.

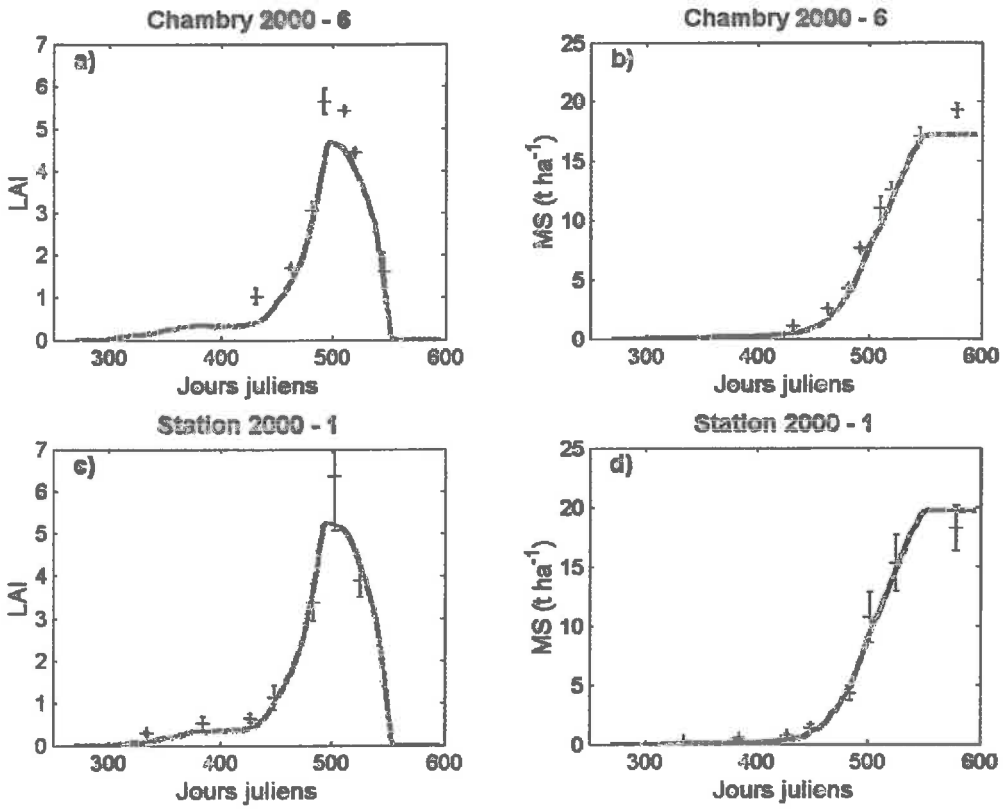


Figure II.4-1 : Evolution des variables *LAI* et la matière sèche observées et simulées sur deux exemples : a) *LAI* simulé (trait continu) et mesuré (croix) par planimétrie sur le traitement 300 kg ha^{-1} de Chambry en 2000 ; b) Matière sèche simulée et observée sur le même site ; c) *LAI* simulé et mesuré par planimétrie sur une station de Chambry en 2000 ; d) Matière sèche simulée et observée sur le même site. Les barres d'erreur indiquent les écarts-types.

Dans le cas de Chambry 2000 - 6, la cinétique de *LAI* mesuré par planimétrie (a) est assez mal reproduite : *LAI* est sous-estimé en début de culture, puis à son maximum. Mais les conséquences sur la matière sèche (b) sont assez minimes : la faiblesse initiale de *LAI* entraîne une sous-estimation de la croissance initiale, puis le retard se rattrape par la suite. Les valeurs maximales de *LAI* correspondent aux meilleures simulations de cette variable. Par contre, la valeur finale de matière sèche qui est capitale pour la détermination du rendement, est sous-estimée. Ceci peut être lié au fait que *LAI* est retombé à 0 de façon trop précoce, stoppant l'accumulation de matière sèche, ce qui ne se voit pas sur le graphe (a). L'absence de mesure

tardive de *LAI* nous empêchant de corriger cette faute, nous avons décidé d'ajuster le paramètre de durée de vie des feuilles, *durvieF*, non par rapport à *LAI* mais par rapport à la matière sèche finale. Le cas de la station (graphes c et d) est presque inverse. *LAI* en phase précoce est assez bien simulé, seul le plateau ne l'est pas. Mais on sait qu'une erreur sur des valeurs de *LAI* supérieures à 3 environ a assez peu d'impact sur la production simulée de matière sèche. Celle-ci est d'ailleurs bien simulée tout au long du cycle. Mais la variable *LAI* retombant à 0 quelque peu trop tard, il s'ensuit une suraccumulation de biomasse en toute fin de cycle.

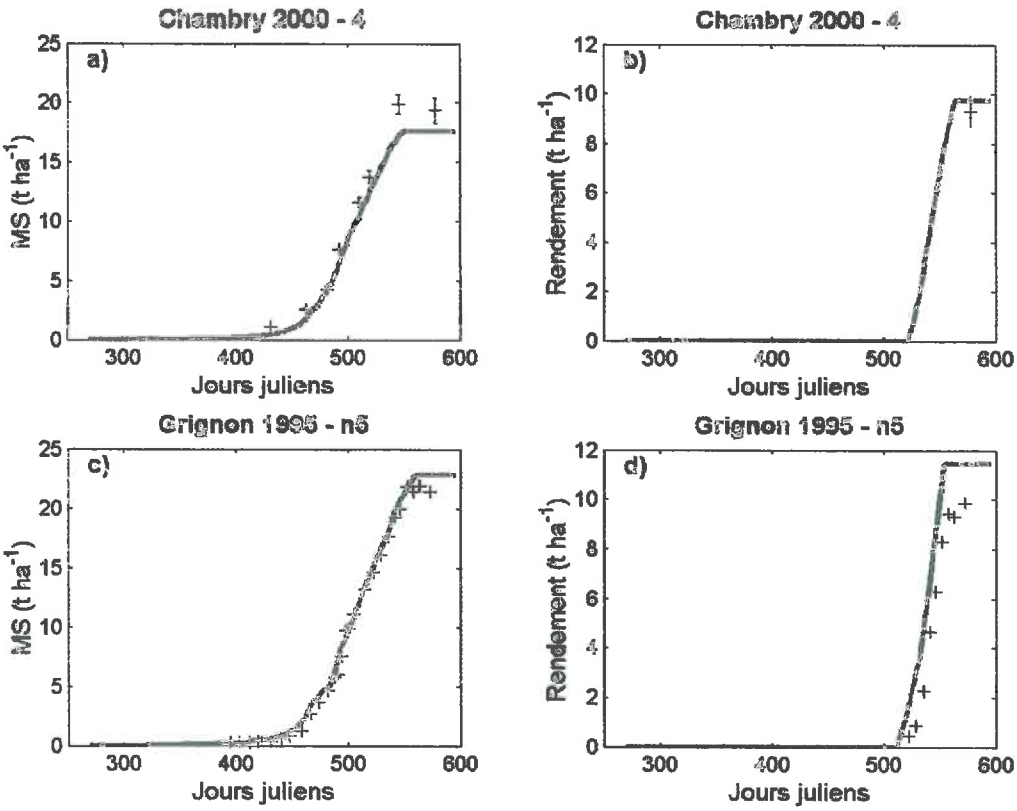


Figure II.4-2 : Evolution de la matière sèche et du rendement observés et simulés sur deux exemples : a) Matière sèche simulée (trait continu) et observée (croix) dans le traitement 220 kg ha^{-1} de Chambry en 2001 ; b) Rendement simulé et observé sur le même essai ; c) Matière sèche simulée et observée dans le traitement de Grignon en 1995 ; d) Rendement simulé et observé sur le même essai. Les barres d'erreur indiquent les écarts-types.

La Figure II.4-2 présente des cinétiques de matière sèche et de rendement en grain. Dans le premier cas (a et b), la matière sèche finale est sous-estimée, ce qui peut être lié à la fin du cycle de *LAI*. Le rendement est par contre surestimé. Ceci peut signifier que la relation matière sèche/rendement est mal paramétrée : en effet, avec une matière sèche sous-estimée, on devrait également avoir un rendement sous-estimé. Par ailleurs, si la matière sèche finale

avait été bien estimée, le rendement aurait été beaucoup trop élevé. Mais des situations inverses se rencontrent, il est donc malaisé de trouver des paramètres qui satisfassent toutes les situations. Le cas de Grignon (c et d) montre une situation où le rendement est trop fort par rapport à la matière sèche finale. Il est surestimé dès le départ, mais l'écart s'accroît en fin de cycle. Ceci est lié à deux effets : la surestimation de l'indice de récolte pour cette situation et la durée trop importante de l'accumulation de biomasse totale et de grains.

La figure suivante présente les résultats de façon plus globale sur *LAI*, la matière sèche et le rendement.

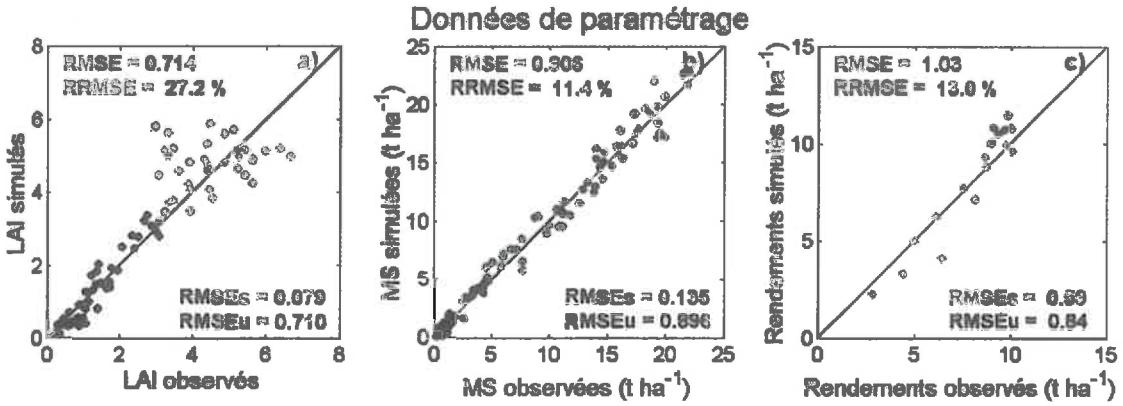


Figure II.4-3 : Variables simulées et observées obtenues sur les données ayant servi à réaliser le paramétrage : a) *LAI* ; b) Matière sèche ; c) Rendement.

LAI est assez mal estimé, avec 29 % d'erreur. Ceci est surtout dû à la grande dispersion qui existe sur les fortes valeurs. L'erreur commise sur les fortes valeurs de *LAI* a peu d'impact lors de la phase de développement foliaire maximal, mais influence la durée de la période pendant laquelle *LAI* est suffisamment important pour produire de la matière sèche. *LAI* est par contre assez peu biaisé, c'est surtout la dispersion qui explique l'erreur commise (0.08 d'erreur systématique contre 0.74 d'erreur non systématique). Les causes de ce genre d'erreurs sont beaucoup plus difficiles à appréhender que celles se traduisant par un biais (Wallach, 2002) et ne sont guère du ressort d'un paramétrage.

La matière sèche est simulée de façon relativement satisfaisante avec seulement 11% d'erreur. Ici aussi, la plus grosse partie de l'erreur commise est liée à la dispersion. Il n'existe cependant pas de lien direct entre les erreurs commises sur *LAI* et celles commises sur la matière sèche (coefficient de corrélation entre les valeurs absolues des écarts *LAI* et de matière sèche : 0.07). Ce résultat est du même ordre de grandeur, bien qu'un peu meilleur,

que celui rapporté par Brisson *et al.* (2002) : 16%. Mais cette dernière valeur est surtout à comparer avec la valeur obtenue dans notre cas en validation.

Les simulations de rendement sont moins bonnes que celles de matière sèche. On commet cependant seulement 13% d'erreur contre les 18% déjà trouvés avec STICS en validation (*cf.* II.3.1.6). Ici encore, cette dernière valeur sera à comparer avec celle obtenue lors de la phase de validation. Pour le rendement, la part de l'erreur systématique est plus forte, montrant que le paramétrage est peut-être moins abouti.

Une part des erreurs observées, comme l'erreur systématique sur le rendement, peut également être liée au formalisme du modèle (erreur de structure). Par ailleurs, il est possible que les valeurs adoptées par les paramètres compensent artificiellement une part de cette erreur de structure.

II.4.2 Validation

Les résultats présentés ici ont été simulés grâce aux variables climatiques observées. Nous comparerons à présent les valeurs simulées et observées des variables d'intérêt, *G*, *P* et *R* et ce d'autant plus que le jeu de données de validation ne contient pas pléthore de valeurs de *LAI* et de matière sèche. De surcroît, les valeurs de *LAI* de La Minière et à La Cheppe ont seulement été obtenues de façon indirecte avec un LICOR. Nous représenterons également la variable *B*, représentant un bilan des entrées et sorties d'azote, définie dans la partie II.2.1.5 et qui sert au calcul du critère agro-environnemental.

La valeur de RRMSE trouvée sur le rendement, *G*, est du même acabit que celle obtenue sur les données de paramétrage avec une légère dégradation : 15.7% contre 13%. On trouve ainsi une valeur légèrement inférieure à celle obtenue avec un jeu de données plus large (*cf.* II.3.1.6), 18.2%. Si on considère les valeurs de RMSEs et de RMSEu, on constate que le rapport RMSEs/RMSEu est du même ordre de grandeur en validation dans les deux cas (ici et dans Brisson *et al.* 2002), 0.21 alors qu'il est beaucoup plus faible pour les données ayant servi au paramétrage : 0.09. Si on est parvenu à diminuer le biais dans les données de paramétrage, il semblerait qu'il augmente lorsqu'on passe aux données de validation. On peut observer que les fortes valeurs de rendement sont assez souvent surestimées. Ceci traduit une tendance à la surestimation de la réponse du rendement à la dose d'azote dans certaines conditions. Nous reviendrons en temps voulu plus longuement sur cette remarque.

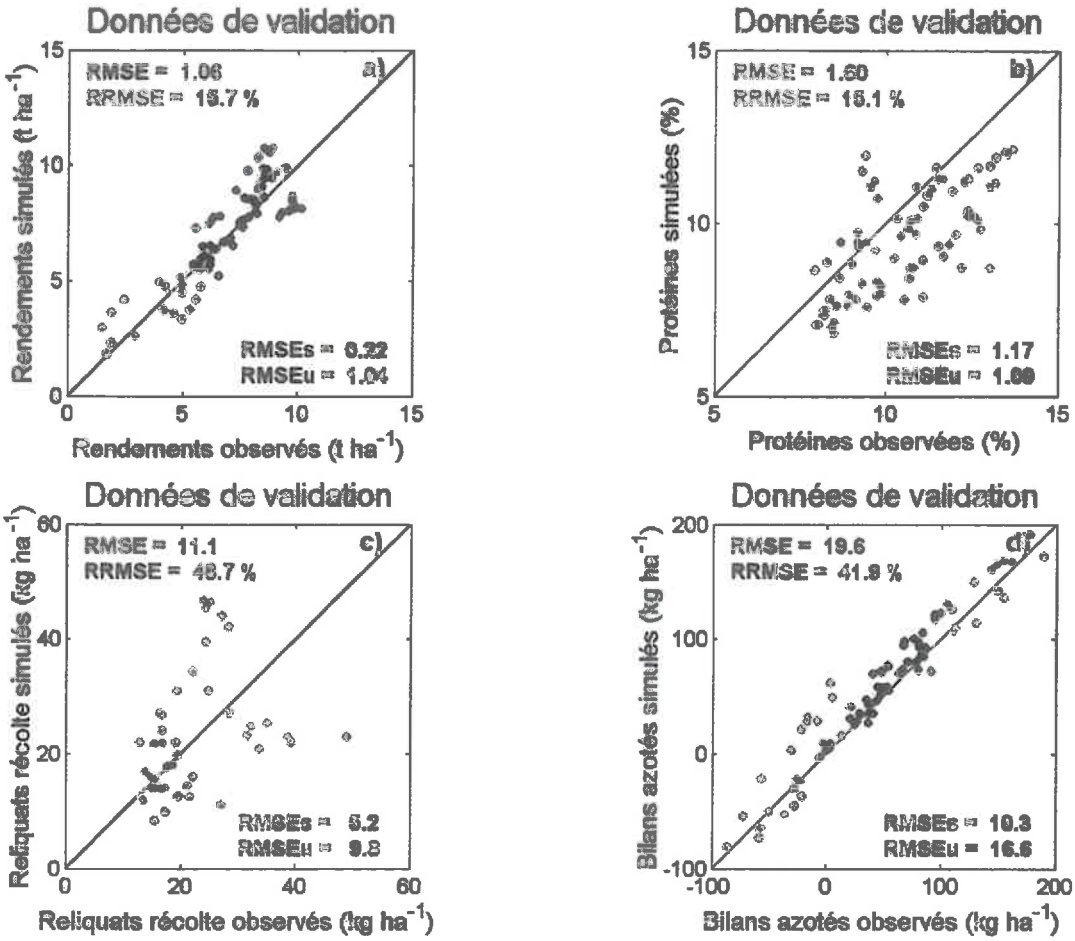


Figure II.4-4 : Valeurs de différentes variables simulées et observées obtenues sur les données ayant servi à réaliser la validation : a) Rendement ; b) Teneur en protéines ; Reliquats d'azote dans le sol à la récolte ; d) Bilan d'azote : entrée moins sortie d'azote sur la parcelle (*cf.* II.2.1.5)

La simulation de la teneur en protéines, P, est en revanche beaucoup plus problématique. On observe une tendance marquée à une sous-estimation systématique. La partie de l'erreur systématique est d'ailleurs plus importante que la partie non-systématique. Une erreur de 1.6% sur la teneur en protéines est relativement importante si l'on pense que le paiement de la qualité des grains se fait dans un intervalle de 2.5% seulement : le prix minimal est en-dessous de 9.5% et le prix maximal, au-dessus de 12%. Mais l'erreur sur cette variable précédemment rapportée dans Brisson *et al.* était de 2.05% de protéines, situant nos résultats dans des gammes acceptées.

La simulation du reliquat d'azote minéral à la récolte, R, est elle aussi affectée de quelques difficultés. Avec une erreur relative de 48% on retrouve le même ordre de grandeur que trouvé par Brisson *et al.* (2002). La valeur beaucoup plus faible de RMSE en valeur absolue est liée au fait que dans notre jeu de données les valeurs observées sont assez faibles (en général inférieures à 40 kg ha^{-1}), ce qui peut expliquer pour une large part la forte erreur

relative. On peut dire que le modèle a simulé de façon satisfaisante le fait que les reliquats à la récolte étaient faibles.

L'autre variable à vocation environnementale, le bilan d'azote, *B*, est nettement mieux simulée que *R*. La RMSE n'est que de l'ordre de 20 kg ha^{-1} pour une gamme de variation allant de –100 à 200 kg ha^{-1}. La valeur importante de la RRMSE est liée au fait qu'on a des valeurs négatives. La part de l'erreur systématique reste relativement importante et est liée à l'erreur d'estimation de *P*. Les bons résultats de la simulation de cette variable sont liés d'une part au fait qu'elle est principalement liée au rendement (qui est lui même bien simulé) et surtout à la dose d'azote qui est supposée parfaitement connue. Cette dernière variable environnementale, de ce point de vue, sera donc plus intéressante à utiliser que le reliquat à la récolte. Les résultats de la simulation sont cependant ici légèrement moins bons que ceux de Brisson *et al.* (2002) qui de plus avait une erreur systématique plus faible (*cf.* II.3.1.6).

Les difficultés de simulation de la teneur en protéines et du reliquat d'azote à la récolte peuvent provenir du fait que les paramètres contrôlant l'absorption de l'azote n'ont pas été modifiés lors de notre paramétrage et que les valeurs établies pour la version précédente du modèle ne sont plus correctes.

II.4.3 Conclusion partielle

Il ressort de ce travail de paramétrage les points suivants :

- les résultats de la version 5 de STICS sont équivalents à ceux décrits par Brisson *et al.* (2002) concernant sa version 4 ;

- la méthode de paramétrage est satisfaisante, notamment par la compréhension du fonctionnement du modèle qu'elle exige, mais il est extrêmement difficile de trouver un jeu de données à la hauteur de ses exigences ; nous avons vu que les informations concernant le sol ont parfois été modifiées au vu des résultats au cours de la validation. Il est probable qu'une meilleure connaissance des variables sol du jeu de données de paramétrage aurait donné des résultats différents. Par ailleurs, le nombre d'essais où toutes les mesures nécessaires au paramétrage ont été faites est insuffisant ;

- les variables *P* et *R* sont relativement mal simulées et il faudra en tenir compte pour l'élaboration du critère agro-environnemental ;

- la variable B est en revanche bien mieux simulée que R ; elle sera donc plus avantageuse à utiliser dans le calcul du critère agro-environnemental ;

- le rendement G est très sensible à un paramètre plante comme *durvieF* ainsi qu'aux entrées sol ; ceci aura son importance dans le cadre de l'assimilation (Chapitre III et IV) ;

- *LAI* est assez mal simulé mais, malgré cela, les rendements le sont assez bien ; ceci est également à retenir pour la partie concernant l'assimilation. En effet, si on corrige le modèle par rapport à une variable qu'il simule mal, les résultats de la correction risquent plus de biaiser le système que de l'améliorer.

A ce stade de ce chapitre, nous avons décrit un certain nombre de critères agro-environnementaux envisageables pour raisonner la fertilisation azotée et les variables qui sont nécessaires à leur calcul. Dans une deuxième phase, nous avons montré la qualité des simulations de ces variables par le modèle de culture STICS. Il s'agit à présent de choisir un critère agro-environnemental sur la base d'un jeu de données et d'évaluer l'aptitude du modèle à réaliser des préconisations à partir de ce critère.

II.5 Choix d'un critère agro-environnemental

Cette partie évalue l'intérêt de quelques critères agro-environnementaux tirés de la bibliographie (II.2.1). Nous nous focaliserons sur les critères de type « fonction objectif » incluant le calcul d'une marge. Les variantes proviendront uniquement de la façon de prendre en compte la contrainte environnementale. Il y aura deux niveaux d'options à croiser :

- la nature de la variable environnementale prise en compte : le reliquat à la récolte (R) ou le bilan d'azote (B) ;

- la façon de l'intégrer dans le calcul du critère : sous forme de taxe ou sous forme de seuil.

Dans cette partie, seules les données expérimentales seront utilisées, afin de pouvoir esquisser un choix parmi des critères sur des données réelles. Il en résulte que, plutôt que de parler de « doses optimales », on devra parler de « meilleure dose parmi celles expérimentées ».

II.5.1 Proposition de différentes formes de critère

La marge brute M se calculera pour tous les critères de la façon suivante :

$$M(N)=G(N)\times f[P(N)]-a.N \qquad \text{Équation II.5-1}$$

où G désigne le rendement en t ha^{-1} à 0% d'humidité, P la teneur en protéines des grains en %, N la dose d'azote utilisée en kg ha^{-1} et a le prix de l'engrais en € kg^{-1}. La fonction f calcule le prix du blé en € t^{-1} en fonction de la teneur en protéines des grains. Elle sera toujours calculée avec les paramètres suivants, correspondant au barème pratiqué en 2002 par la coopérative COHESIS de Laon :

$f[P(N)]=86.25$ si $P(N) \leq 9$

$f[P(N)]=86.25 + 3.93\ (P(N) - 9)$ si $9 < P(N) \leq 12.5$

$f[P(N)]=100$ si $P(N) > 12.5$

Le prix de l'engrais sera celui pratiqué par la même coopérative sur l'ammonitrate, soit :

$a = 0.60$

On rappelle que la prise en compte de P de cette façon lui donne assez peu de poids. C'est un inconvénient du point de vue de l'optimisation de la qualité de la production, mais vu que cette variable est mal simulée, il est bon de ne pas lui attribuer trop d'importance.

Quatre types de critères agro-environnementaux ont été envisagés (Tableau II.5-1). Les deux premiers consistent à utiliser une taxation basée soit sur le bilan d'azote, soit sur les reliquats récolte. La fonction objectif est alors égale à la marge brute diminuée du montant d'une taxe proportionnelle au bilan d'azote (critère A) ou aux reliquats azotés (critère B). Les deux autres imposent quant à eux une valeur limite supérieure au bilan d'azote (critère C) ou au reliquat (critère D). Pour ces deux derniers critères, la dose optimale sera celle qui maximise la marge brute parmi celles qui conduisent à des valeurs de bilans d'azote ou de reliquats inférieures à la valeur limite ou seuil. Si aucune dose ne permet d'être en dessous de ce seuil, on prend celle qui permet de s'en rapprocher le plus.

Critère	Fonction objectif	Contrainte
A	$J(N)=M(N) - \alpha . B(N)$	-
B	$J(N)=M(N) - \beta . R(N)$	-
C	$J(N)=M(N)$	N^* telle que $B(N^*) < \gamma$
D	$J(N)=M(N)$	N^* telle que $R(N^*) < \delta$

Tableau II.5-1 : Critères agro-environnementaux utilisés : A, la fonction objectif est égale à la marge brute diminuée d'une taxe fonction du bilan d'azote selon un coefficient α ; B, la fonction objectif est égale à la marge brute diminuée d'une taxe fonction du reliquat récolte selon un coefficient β ; C, la fonction objectif est égale à la marge brute et une contrainte sur le bilan d'azote est imposée sous forme d'une valeur limite supérieure, γ ; D, la fonction objectif est égale à la marge brute et une contrainte sur le reliquat récolte est imposée sous forme de valeur limite supérieure, δ, où α, β, γ et δ sont positifs.

Dans un premier temps, deux valeurs du niveau de la taxation défini par les paramètres α et β et des valeurs limites γ et δ seront étudiées (II.5.2). Lors de la comparaison des critères A, B, C et D, une étude de leur comportement sera réalisée pour une évolution continue des valeurs de α, β, γ et δ (II.5.3).

Il a été envisagé d'utiliser comme variable environnementale (E) la concentration en nitrates de l'eau drainée au cours de l'hiver suivant la récolte, sous hypothèse d'un sol nu. En effet, cette variable est la plus pertinente du point de vue environnemental pour maîtriser la pollution nitrique. Mais son utilisation pose plusieurs problèmes :

- cette variable est rarement mesurée expérimentalement, en particulier sur des essais de fertilisation azotée : il n'était donc pas possible de tester un critère agro-environnemental l'impliquant sur un jeu de données expérimentales ;

- une première série de tests a montré que le modèle simulait un faible impact de la dose d'azote sur cette variable ; en outre, son utilisation doublait le temps de calcul.

Nous n'avons donc considéré que le bilan d'azote et les reliquats d'azote à la récolte.

II.5.2 Application de ces critères pour deux niveaux de taxation ou de seuil

Les données utilisées pour tester ces différents critères agro-environnementaux correspondent aux essais de doses d'azote présentés dans le Tableau II.3-6. Ceci représente, avec les trois essais de Boigneville 1998 et 1999, un total de 14 essais.

II.5.2.1 Valeurs de α, β, γ et δ utilisées

Les valeurs utilisées dans cette partie pour les différents critères sont données dans le Tableau II.5-2.

Critère	Hypothèse	Valeur du paramètre
A	Basse	$\alpha = 0.23$ € kg^{-1}
A	Haute	$\alpha = 1.50$ € kg^{-1}
B	Basse	$\beta = 1.00$ € kg^{-1}
B	Haute	$\beta = 6.00$ € kg^{-1}
C	Basse	$\gamma = 90$ kg ha^{-1}
C	Haute	$\gamma = 60$ kg ha^{-1}
D	Basse	$\delta = 35$ kg ha^{-1}
D	Haute	$\delta = 20$ kg ha^{-1}

Tableau II.5-2 : Valeurs de α, β, γ et δ retenues respectivement pour les critères A, B, C et D. L'hypothèse basse correspond à un scénario de faible contrainte environnementale ; l'hypothèse haute à un scénario de plus forte contrainte.

La valeur de α=0.23 € kg^{-1} pour le critère A correspond à la valeur de taxe pressentie dans le projet de loi ajourné (II.2.1.5). Cette valeur est relativement faible, ce qui apparente le critère A dans cette configuration à la situation de rémunération actuelle. La valeur de δ sous

hypothèse basse provient de Makowski *et al.* (2001). En première approche, on suppose que le reliquat est mesuré sur 90 cm. Les hypothèses haute et basse pour les différents critères ne sont naturellement pas directement comparables, mais elles sont censées traduire des niveaux de préoccupation environnementale plus ou moins importants.

II.5.2.2 Réponse des variables utilisées à la dose d'azote et premières conséquences

La Figure II.5-1 illustre à l'aide de six essais la diversité des courbes de réponses à la fertilisation azotée obtenue sur les quatorze essais utilisés. Elle représente les différentes variables impliquées dans le calcul des critères agro-environnementaux, *G*, *R*, *B*, *M* et la valeur du critère agro-environnemental A, pour les hypothèses basse et haute. La large gamme de rendements obtenue (de 1.5 à 9.9 t ha^{-1} à 0% d'humidité) est due tant à la variabilité des doses d'azote employées qu'à la diversité des conditions pédologiques rencontrées. La forme de ces courbes est classique (*e.g.* Makowski *et al.*, 2001) : un rendement maximum est clairement atteint dans certains cas (essais 1 et 4), mais pas dans d'autres (essais 8 et 13).

Les reliquats azotés à la récolte ont des réponses similaires à ce qui a été décrit dans la bibliographie (Chaney, 1990 ; Makowski *et al.*, 2001). Lorsque la dose d'azote augmente, le reliquat reste d'abord stable avant d'augmenter, surtout lorsque la dose optimale du point de vue du rendement est dépassée. L'essai 8 est un contre-exemple de cette règle puisqu'on observe une augmentation assez régulière depuis la dose 0. Le bilan d'azote augmente linéairement en fonction de la dose d'azote dans tous les essais. L'ordonnée à l'origine varie entre –72 et –21 kg ha^{-1} mais la pente ne varie guère : 0.55 ± 0.07. Ceci est cohérent avec les références citées par Mary *et al.* (2002). On peut remarquer que le bilan d'azote obtenu avec la dose donnant l'optimum de rendement est toujours positif : de 17 à 120 kg ha^{-1}. Cette même figure montre en outre les évolutions de la marge brute obtenue pour ces mêmes essais ainsi que celle du critère A pour les hypothèses haute et basse (voir Tableau II.5-2). On vérifie que le critère A, avec une faible taxation environnementale, ressemble de près à la marge brute actuelle : la contrainte environnementale exercée dans ce cas est donc très faible. Une taxation plus forte, comme illustrée avec le cas de l'hypothèse haute, modifie la forme de la réponse et donc la valeur de la dose d'azote optimale dans la plupart des essais, à quelques exceptions près (essai 1).

La partie suivante présente de façon plus exhaustive l'évolution des meilleures doses d'azote pour chacun des essais et les conséquences induites sur les variables *G*, *P*, *R* et *B*.

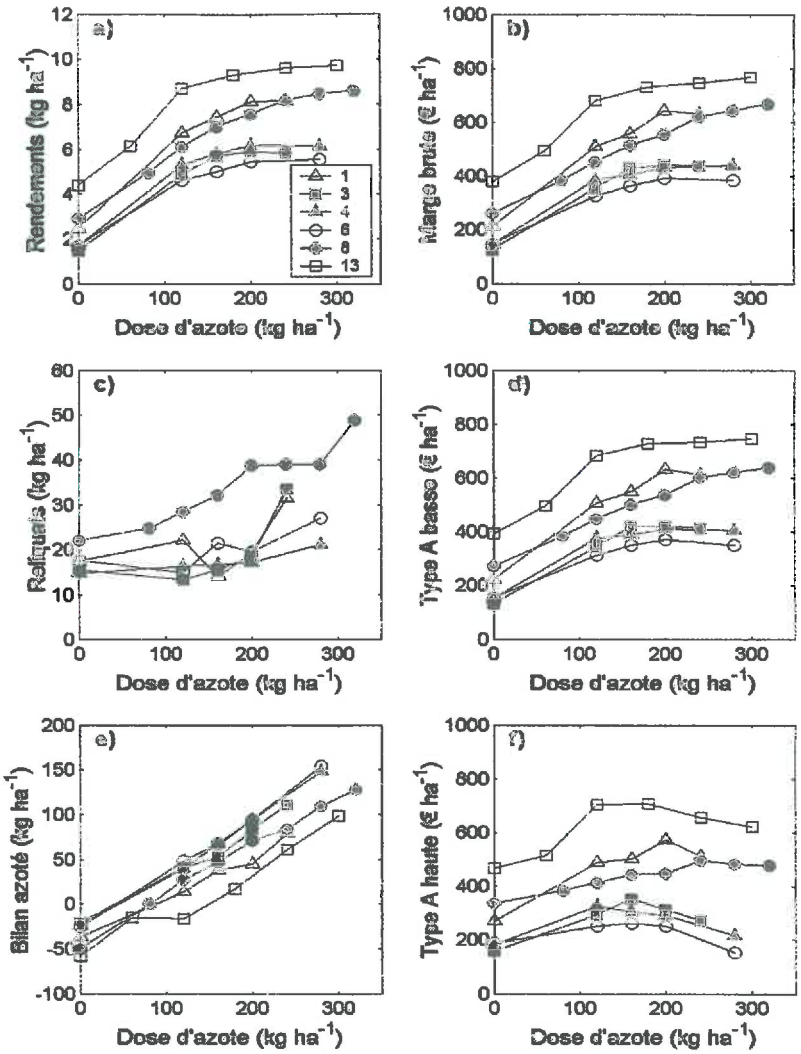


Figure II.5-1 : Courbes de réponse observées sur 6 essais : a) Rendement ; b) Marge brute ; c) Reliquats azotés à la récolte ; d) Fonction objectif A avec α=0.23 € kg^{-1} ; e) Bilan d'azote à la récolte ; f) Fonction objectif A avec α=1.50 € kg^{-1}. Les numéros des essais renvoient au Tableau II.3-6.

II.5.2.3 Détermination des meilleurs doses pour les différentes valeurs de paramètres

La meilleure dose obtenue pour chaque essai et chaque type de critère pour les hypothèses haute et basse est donnée dans le Tableau II.5-3. Dans la plupart des situations (12 essais sur 14), elle dépend du critère choisi et de l'hypothèse faite sur la valeur du paramètre. Par exemple, la dose optimale varie fortement dans l'essai 4, de 120 à 280 kg ha^{-1}, alors qu'elle reste la même pour les essais 1 et 12. Ces différences de comportement résultent de la diversité des courbes de réponse : par exemple, dans le cas de l'essai 4, le rendement et les reliquats varient très peu en fonction de la dose d'azote. En conséquence, une variation de la

contrainte environnementale bouleverse l'ordre des doses d'azote. Par contre, les réponses des variables de l'essai 1 varient beaucoup plus fortement, ce qui stabilise la hiérarchie entre les résultats des différentes doses lorsqu'on change de critère de sélection de la dose optimale. Le fait d'utiliser un critère à seuil peut conduire à recommander des doses nulles : c'est le cas pour le critère D, hypothèse haute concernant les essais 8 et 12. Dans ces essais, les reliquats étaient en effet toujours supérieurs au seuil de 20 kg ha^{-1}. Cet exemple montre la difficulté d'utiliser un tel critère avec un seuil aussi sévère et identique pour tous les types de sol. L'hypothèse basse utilisée avec le même type de critère (D-b) donne des résultats plus cohérents. Ceci illustre en outre que l'utilisation du reliquat dans le critère peut être difficile à généraliser pour deux raisons principales : (i) il varie peu en fonction de la dose d'azote, si bien que des erreurs d'évaluation peuvent avoir des conséquences importantes sur les préconisations de fertilisation azotée et (ii) il est fortement influencé par le type de sol (par exemple, un sol argileux laisse un reliquat d'azote plus important qu'un sol limoneux ou sableux, voir notamment Booltink *et al.*, 2001), si bien qu'il faudrait que le seuil varie en fonction de la zone pédologique.

Le Tableau II.5-4 présente, pour chaque type de critère proposé et chaque valeur de paramètre testé, les résultats obtenus à la meilleure dose en terme de valeur de la fonction objectif (incluant ou non une taxe selon les types de critères). Le Tableau II.5-5 donne les valeurs de bilan d'azote et le Tableau II.5-6 indique les valeurs de rendement. On constate que, si on compare pour un même type de critère les hypothèses haute et basse de contrainte environnementale, les doses d'azote recommandées sont effectivement plus faibles dans le cas de l'hypothèse basse. En moyenne, on observe une baisse de 48, 20, 40 et 67 kg ha^{-1} pour les critères A, B, C et D respectivement. Ces valeurs ne sont pas directement comparables entre elles, vu que les valeurs de paramètres choisies sont arbitraires. Les pertes de revenu correspondantes sont de 89, 108, 38 et 77 € ha^{-1} (Tableau II.5-4). La baisse du bilan d'azote est de 30, 13, 22 et 38 kg ha^{-1} respectivement (Tableau II.5-5) et les pertes de rendement sont de 0.4, 0.0, 0.5 et 1.1 t ha^{-1} (Tableau II.5-6).

Les types de critères A et B, basés sur une taxation, sont doublement pénalisants : si l'on compare en effet les différences entre hypothèses haute et basse des critères A et C, on constate que, pour une amélioration du bilan environnemental équivalente (respectivement de 30 et 22 kg ha^{-1}), la baisse de revenu est beaucoup moins forte dans le cas du critère C que du critère A (38 contre 89 € ha^{-1}). Ceci provient de ce que dans le critère A, la baisse de revenu est due à deux causes : la diminution du rendement à la meilleure dose et la taxation. Une

comparaison similaire effectuées sur les critères C et D conduit à une conclusion équivalente : avec un seuil (type D), la perte entre hypothèse haute et basse est plus faible qu'avec une taxation (type B) pour les valeurs de paramètres choisies (77 et 108 € ha^{-1} respectivement pour D et B). Mais l'amélioration du bilan environnemental est bien meilleure pour D (38 contre 13 kg ha^{-1}). Ainsi, il apparaît assez clairement sur ces exemples que des critères basés sur un seuil peuvent avoir les mêmes performances sur le plan environnemental que des critères basés sur des taxes mais avec moins de répercussions sur le revenu de l'agriculteur, voire sur la production.

Il semble donc que les meilleurs critères soient ceux basés sur un seuil du bilan d'azote. Ils sont plus efficaces du point de vue environnemental et diminuent moins le revenu de l'agriculteur par l'effet de la taxe. Mais avec l'approche utilisée ici, on ne peut pas comparer les performances économiques des différents critères à performance environnementale équivalente. Pour ce faire, il est nécessaire d'adopter une approche consistant à faire varier de façon continue les valeurs de α, β, γ et δ, ce qui est précisément l'objet de la partie suivante.

Essais	Meilleure dose (kg ha^{-1})							
	A-b	A-h	B-b	B-h	C-b	C-h	D-b	D-h
1	200	200	200	200	200	200	200	200
2	240	240	240	240	240	160	240	240
3	200	160	200	160	200	160	200	200
4	200	120	280	200	160	120	280	200
5	220	220	220	220	180	140	300	220
6	200	160	200	200	160	120	200	200
7	240	240	-	-	240	160	-	-
8	320	240	320	280	240	160	160	0
9	300	150	-	-	250	200	-	-
10	300	200	-	-	250	200	-	-
11	190	190	-	-	190	190	-	-
12	220	220	220	220	220	220	220	0
13	300	180	-	-	240	180	-	-
14	280	220	-	-	120	120	-	-
Moyenne[(1)]	244	196			206	166		
Moyenne [(2)]	225	195	235	215	200	160	225	158
Différence b-h	48/30		20		40/40		67	

Tableau II.5-3 : Valeurs des meilleures doses pour les critères A, B, C et D pour les hypothèses basse (b) ou haute (h) de la valeur des paramètres α, β, γ et δ. [(1)] Moyenne des 14 essais ; [(2)] Moyenne des 8 essais pour lesquels toutes les variables ont été mesurées. Différence b-h : différence entre la moyenne obtenue avec l'hypothèse basse et l'hypothèse haute, calculée d'abord le cas échéant avec la moyenne des 14 essais puis avec celle des 8 essais.

Essais	Valeurs du critère (€ ha^{-1})							
	A-b	A-h	B-b	B-h	C-b	C-h	D-b	D-h
1	631	574	622	525	641	641	641	641
2	581	478	580	483	599	521	599	599
3	420	353	421	340	440	433	440	440
4	410	325	419	329	406	386	440	432
5	440	314	450	386	430	350	467	463
6	372	262	374	276	364	325	393	393
7	579	470	-	-	599	486	-	-
8	640	496	620	410	622	513	513	261
9	643	523	-	-	645	602	-	-
10	679	565	-	-	666	635	-	-
11	868	889	-	-	864	864	-	-
12	706	662	690	565	715	715	715	364
13	745	706	-	-	747	731	-	-
14	577	422	-	-	481	481	-	-
Moyenne[(1)]	592	503			587	549		
Moyenne [(2)]	525	433	522	414	527	486	526	449
Différence b-h	89/92		108		38/41		77	

Tableau II.5-4 : Valeurs des critères A, B, C et D calculés aux meilleures doses pour les hypothèses basse (b) et haute (h) de la valeur des paramètres α, β, γ et δ. [1] Moyenne des 14 essais ; [2] Moyenne des 8 essais pour lesquels toutes les variables ont été mesurées. Différence b-h : différence entre la moyenne obtenue avec l'hypothèse basse et l'hypothèse haute, calculée d'abord le cas échéant avec la moyenne des 14 essais puis avec celle des 8 essais.

Essais	Bilan d'azote ($kg\ ha^{-1}$)							
	A-b	A-h	B-b	B-h	C-b	C-h	D-b	D-h
1	45	45	45	45	45	45	45	45
2	81	81	81	81	81	44	81	81
3	84	53	84	53	84	53	84	84
4	94	40	148	94	67	40	148	94
5	100	100	100	100	80	60	161	100
6	94	68	94	94	68	48	94	94
7	86	86	-	-	86	49	-	-
8	128	84	128	109	84	47	47	-50
9	112	25	-	-	82	54	-	-
10	105	47	-	-	76	47	-	-
11	-17	-17	-	-	-17	-17	-	-
12	35	35	35	35	35	35	35	-59
13	98	17	-	-	61	17	-	-
14	130	97	-	-	40	40	-	-
Moyenne[(1)]	84	54			62	40		
Moyenne [(2)]	83	63	89	76	68	47	87	49
Différence b-h	30/20		13		22/21		38	

Tableau II.5-5 : Valeurs des bilans azotés pour les critères A, B, C et D obtenus aux meilleures doses pour les hypothèses basse (b) et haute (h) de la valeur des paramètres α, β, γ et δ. [(1)] Moyenne des 14 essais ; [(2)] Moyenne des 8 essais pour lesquels toutes les variables ont été mesurées. Différence b-h : différence entre la moyenne obtenue avec l'hypothèse basse et l'hypothèse haute, calculée d'abord le cas échéant avec la moyenne des 14 essais puis avec celle des 8 essais.

Essais	Rendement ($t\ ha^{-1}$)							
	A-b	A-h	B-b	B-h	C-b	C-h	D-b	D-h
1	8.1	8.1	8.1	8.1	8.1	8.1	8.1	8.1
2	7.6	7.6	7.6	7.6	7.6	7.0	7.6	7.6
3	5.9	5.7	5.9	5.7	5.9	5.7	5.9	5.9
4	6.2	5.3	6.2	6.2	5.8	5.3	6.2	6.2
5	6.4	6.4	6.4	6.4	6.2	5.0	6.6	6.4
6	5.4	5.0	5.4	5.4	5.0	4.6	5.4	5.4
7	7.8	7.8	-	-	7.8	6.6	-	-
8	8.6	8.2	8.6	8.5	8.2	7.0	7.0	2.9
9	8.5	7.3	-	-	8.2	7.8	-	-
10	8.8	8.1	-	-	8.4	8.1	-	-
11	10.1	10.1	-	-	10.1	10.1	-	-
12	8.5	8.5	8.5	8.5	8.5	8.5	8.5	4.2
13	9.7	9.3	-	-	9.6	9.3	-	-
14	8.6	8.1	-	-	6.4	6.4	-	-
Moyenne[(1)]	7.9	7.5			7.6	7.1		
Moyenne [(2)]	7.1	6.9	7.1	7.1	6.9	6.4	6.9	5.8
Différence b-h	0.4/0.2		0.0		0.5/0.5		1.1	

Tableau II.5-6 : Valeurs des rendements pour les critères A, B, C et D obtenus aux meilleures doses pour les hypothèses basse (b) et haute (h) de la valeur des paramètres α, β, γ et δ. [1] Moyenne des 14 essais ; [2] Moyenne des 8 essais pour lesquels toutes les variables ont été mesurées. Différence b-h : différence entre la moyenne obtenue avec l'hypothèse basse et l'hypothèse haute, calculée d'abord le cas échéant avec la moyenne des 14 essais puis avec celle des 8 essais.

II.5.3 Application de ces critères pour une évolution continue des valeurs de taxation et de seuil

Dans cette partie, la variation des différents critères est examinée en fonction d'une variation continue des paramètres α, β, γ et δ. Plus précisément, une assez large gamme de variation possible pour ces valeurs a été choisie et discrétisée à un pas très fin. Les calculs développés dans la partie précédente ont été effectués pour tous les pas de l'intervalle.

La Figure II.5-2 concerne le critère A. Lorsque la valeur de la taxation augmente, la dose préconisée moyenne (b) diminue régulièrement. Ceci a pour conséquence directe une baisse du rendement (a) et donc de la marge brute sans application de la taxe. Le bilan d'azote diminue également jusqu'à devenir négatif (c). Ceci a pour conséquence de faire réaugmenter la marge incluant la taxe (d) à partir du moment où *B* devient négatif, car la taxe devient alors positive. Cela revient, en quelque sorte, à subventionner les pratiques agricoles à faibles intrants d'azote. Mais d'après notre jeu de données, ce changement intervient pour une valeur de α de 2.7 € kg^{-1}, soit presque 12 fois la valeur prévue dans le projet de loi.

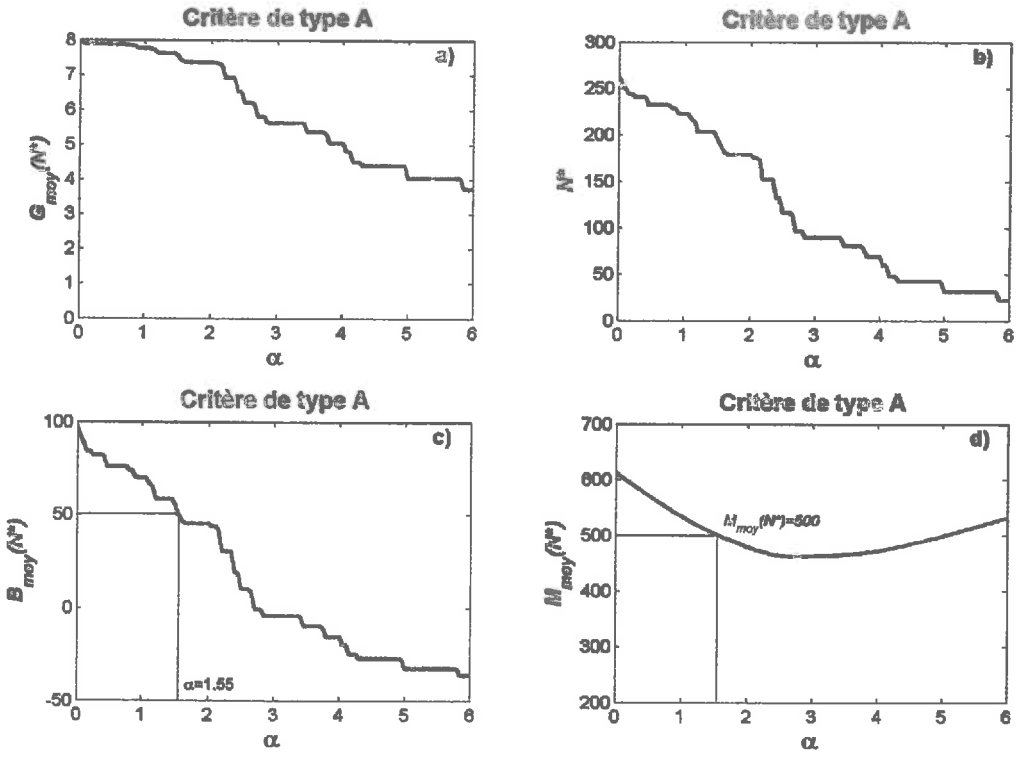


Figure II.5-2 : Evolution de plusieurs variables en fonction du paramètre α du critère A ; valeurs moyennes des 14 essais obtenues avec la meilleure dose ; a) Rendement moyen (t ha^{-1}) ; b) Moyenne des meilleures doses (kg ha^{-1}) ; c) Bilan d'azote moyen (kg ha^{-1}) ; d) Marge brute moyenne incluant la taxation (€ ha^{-1}) (valeur du critère). La valeur de α permettant d'obtenir un bilan d'azote moyen de 50 kg ha^{-1} est mise en évidence.

La Figure II.5-3 donne un aperçu du comportement du critère B en fonction de la valeur de β. On constate que la faible variation de dose optimale et de rendement observée pour deux valeurs de β se retrouve ici pour une large gamme de variation. Ceci est dû à la très faible variation des reliquats d'azote en fonction de la dose. Les doses fortes qui conduisent à des reliquats importants ne sont pas éliminées pour les faibles valeurs de β peu pénalisantes. Puis, lorsque β augmente, elles sont systématiquement éliminées. Les doses plus faibles conduisant à des reliquats proches entre elles, la taxation affecte de la même façon toutes les doses : une augmentation de cette taxation a dès lors peu d'influence. Mais comme les doses élevées conduisent et à des rendements maximaux, l'évolution de B suit celle de N^*. Quant à la valeur de la marge incluant la taxe, elle est quasiment strictement linéaire en fonction de β, car presque rien d'autre ne l'influence. Ceci confirme l'idée que ce critère n'est guère intéressant pour préconiser des doses d'azote.

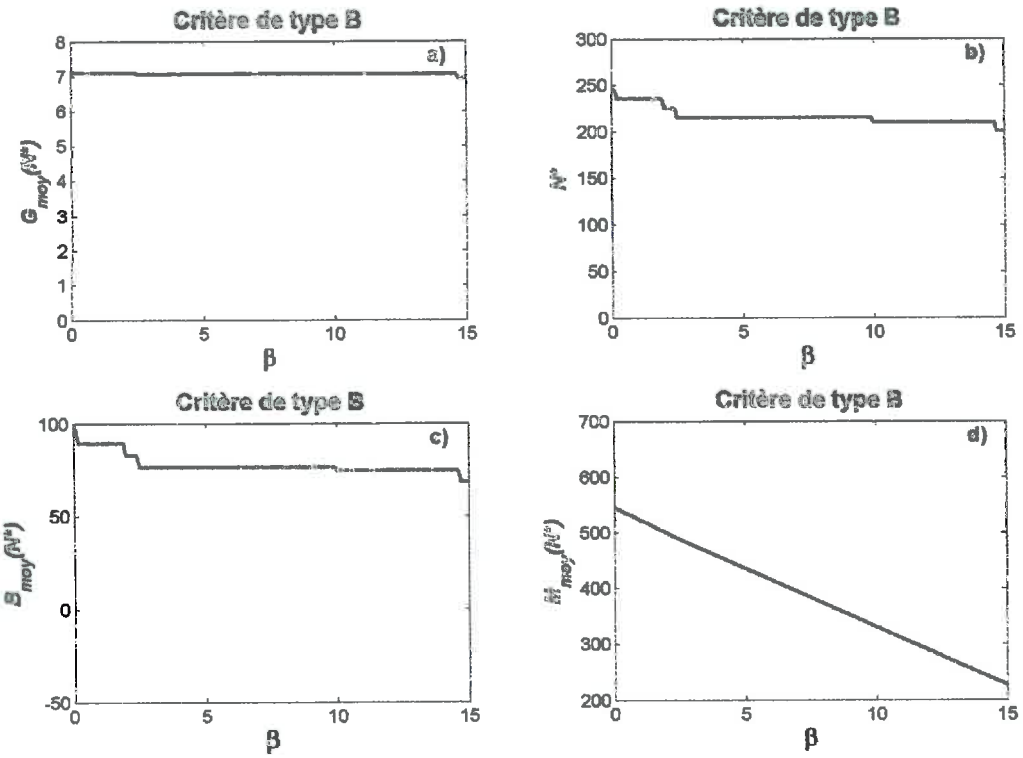


Figure II.5-3 : Evolution de plusieurs variables en fonction du paramètre β du critère B ; valeurs moyennes des 14 essais obtenues avec la meilleure dose ; a) Rendement moyen (t ha^{-1}) ; b) Moyenne des meilleures doses (kg ha^{-1}) ; c) Bilan d'azote moyen (kg ha^{-1}) ; d) Marge brute moyenne incluant la taxation (€ ha^{-1}) (valeur du critère).

La Figure II.5-4 permet d'examiner quelques propriétés du critère C. L'axe des abscisses est inversé de façon à ce que la contrainte environnementale augmente toujours lorsqu'on se déplace vers la droite. Bien que les doses d'azote préconisées commencent à

diminuer sensiblement de façon assez précoce, les rendements obtenus ne sont affectés que beaucoup plus tardivement, vers une valeur de γ de 75 kg ha^{-1} environ. De ce fait, le bilan d'azote diminue également rapidement avec l'augmentation de la contrainte, alors que la marge dégagée ne commence à vraiment diminuer qu'à partir d'une valeur de γ comprise entre 75 et 100 kg ha^{-1}.

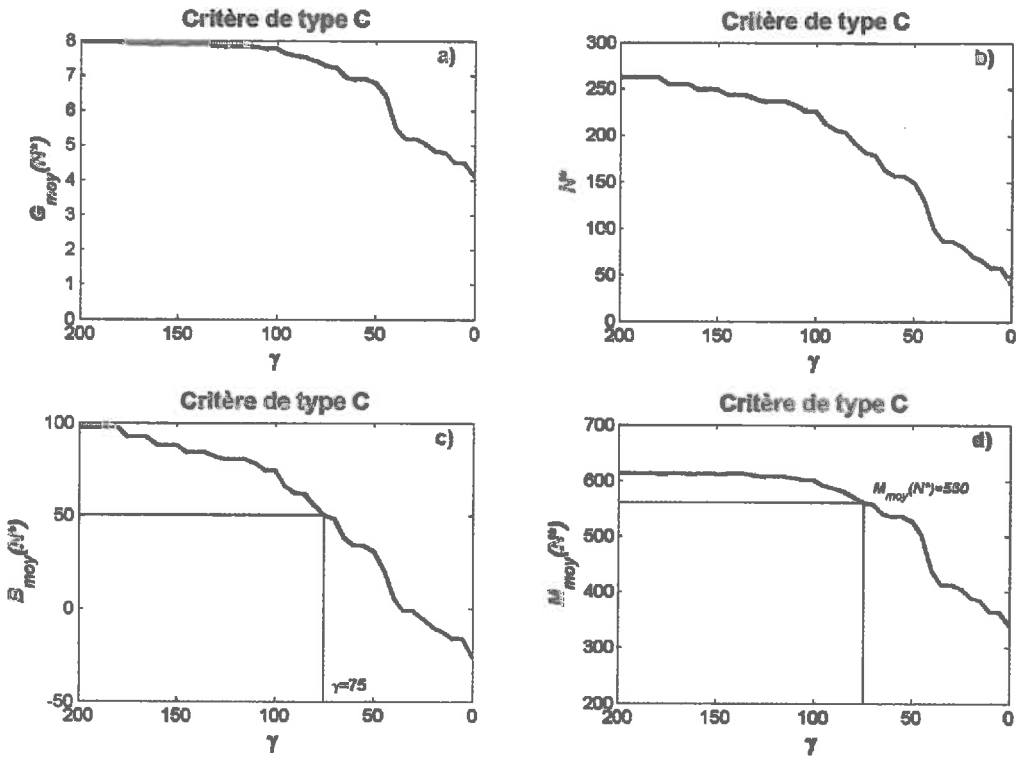


Figure II.5-4 : Evolution de plusieurs variables en fonction du paramètre γ du critère C ; valeurs moyennes des 14 essais obtenues avec la meilleure dose ; a) Rendement moyen (t ha^{-1}) ; b) Moyenne des meilleures doses (kg ha^{-1}) ; c) Bilan d'azote moyen (kg ha^{-1}) ; d) Marge brute moyenne (€ ha^{-1}). L'axe des abscisses est inversé, la contrainte environnementale augmente lorsqu'on va vers la droite. La valeur de γ permettant d'obtenir un bilan d'azote moyen de 50 kg ha^{-1} est mise en évidence.

La Figure II.5-5 représente l'évolution du critère D en fonction de la variation de δ. L'axe des abscisses est ici aussi inversé. On a vu que les reliquats à la récolte étaient assez peu variables pour la plupart des doses d'azote. Ceci a pour effet que, pour un essai donné, on aura un comportement de tout ou rien : au-dessus d'une certaine valeur de δ, presque toutes les doses permettront de satisfaire la contrainte environnementale tandis qu'en-dessous, plus aucune ne le permettra, ce qui conduira à choisir la dose minimale (plus exactement, celle qui permet de se rapprocher le plus du seuil, donc souvent la dose minimale). On a déjà observé ce phénomène dans le Tableau II.5-3 : lorsqu'on passait à une valeur de δ de 20 kg ha^{-1}, deux

essais avaient une dose recommandée nulle. Si l'on continue à baisser le seuil δ, la dose optimale devient nulle dans un nombre croissant d'essais, ce qui fait baisser la moyenne.

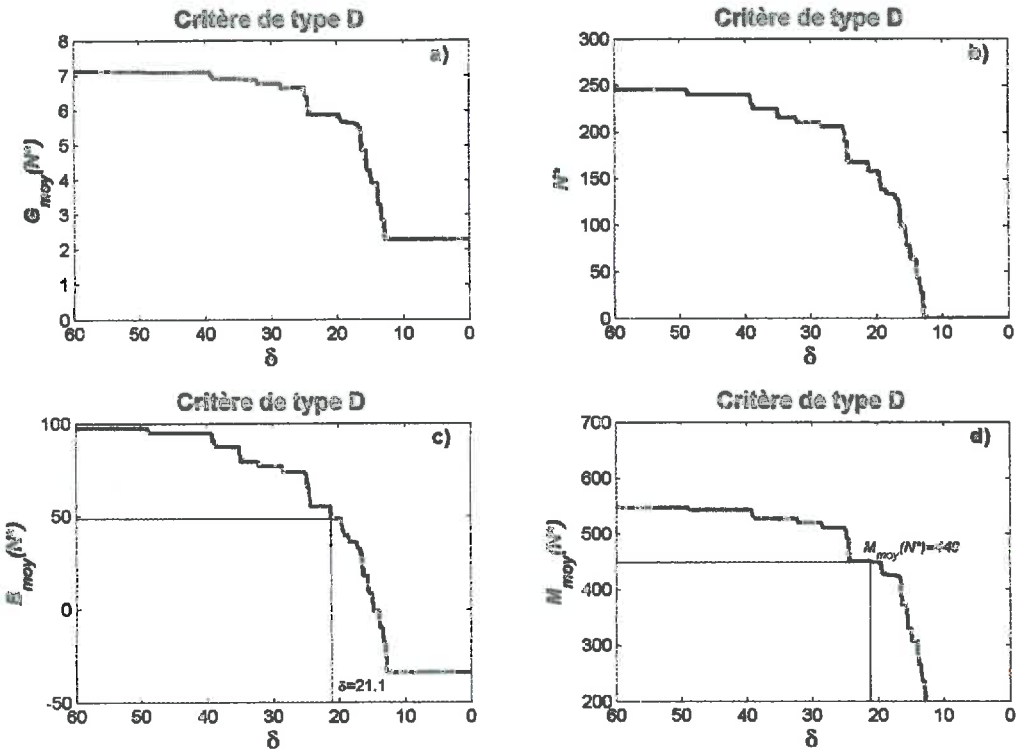


Figure II.5-5 : Evolution de plusieurs variables en fonction du paramètre δ du critère D ; valeurs moyennes des 14 essais obtenues avec la meilleure dose ; a) Rendement moyen (t ha^{-1}) ; b) Moyenne des meilleures doses (kg ha^{-1}) ; c) Bilan d'azote moyen (kg ha^{-1}) ; d) Marge brute moyenne (€ ha^{-1}). L'axe des abscisses est inversé, la contrainte environnementale augmente lorsqu'on va vers la droite. La valeur de δ permettant d'obtenir un bilan d'azote moyen de 50 kg ha^{-1} est mise en évidence.

Ce qui précède est explicité sur un exemple théorique dans la Figure II.5-6. Dans le premier cas (a et d), le seuil est au-dessus des reliquats obtenus avec toutes les doses : la meilleure dose sera donc la plus forte, car elle maximise la marge. Dans le deuxième (b et e), le seuil élimine la plus forte dose : la meilleure dose sera donc la suivante car elle maximise la marge parmi les doses restantes. Dans le troisième (c et f), toutes les doses sont éliminées d'un coup : la meilleure dose sera donc la dose minimale car elle se rapproche le plus du seuil. Pour une faible variation de seuil, la dose optimale est passée très brutalement d'une valeur forte à une valeur faible.

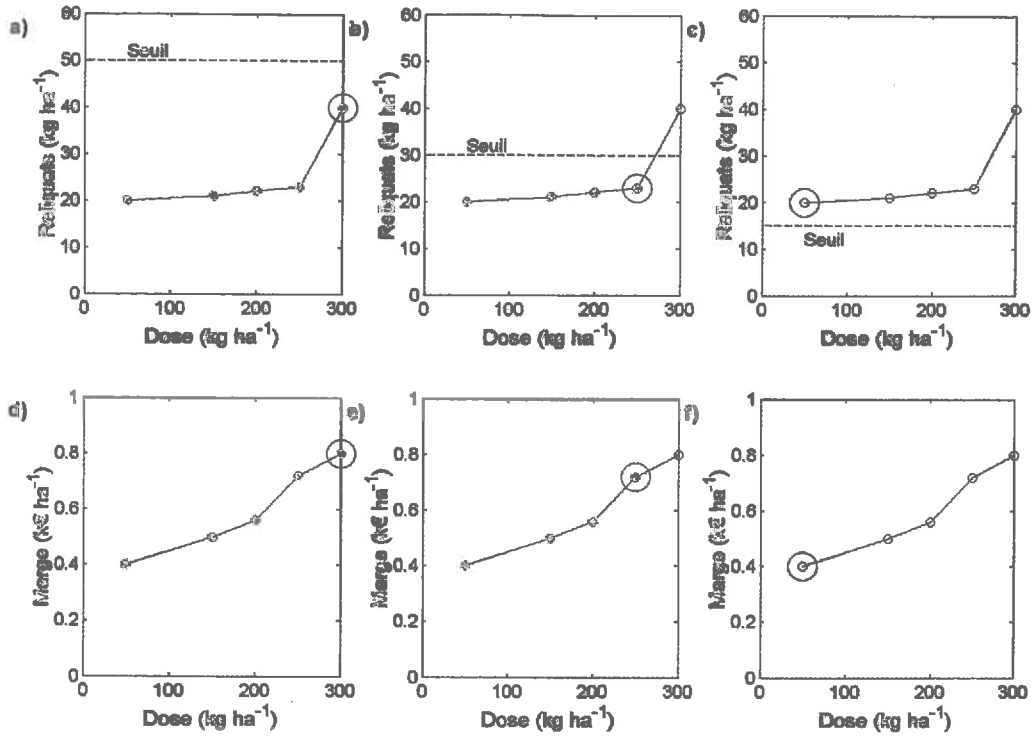


Figure II.5-6 : Illustration sur un cas théorique de l'effet brutal de l'évolution du seuil sur le reliquat ; a) et d) Toutes les doses sont sélectionnées après application du seuil, la meilleure dose est la plus forte ; b) et e) Seule la plus forte dose est rejetée par l'application du seuil, la deuxième est la meilleure dose ; c) et f) Toutes les doses sont rejetées après application du seuil, la plus faible est la meilleure dose.

Cette représentation continue de l'évolution de G, N^*, B et M en fonction des valeurs de α, β, γ et δ permet de comparer les différents critères. La valeur du paramètre permettant d'atteindre un bilan d'azote de 50 kg ha^{-1} a été matérialisé sur toutes les figures présentant l'évolution des différents critères, sauf dans le cas du critère B où ce n'est pas possible. Pour cette valeur de paramètre, la marge brute obtenue est de 500 € ha^{-1} pour le critère A, de 560 € ha^{-1} pour le critère C et de 449 € ha^{-1} pour le critère D. C'est donc le critère C qui permet d'obtenir le meilleur revenu agricole pour un résultat donné du bilan d'azote. Il faut cependant remarquer que le critère D n'a pas été construit avec le même jeu de données, puisque seuls les essais comportant des mesures de reliquat récolte ont pu être utilisés dans son cas. Cette façon de voir les choses est généralisée dans la Figure II.5-7. Cependant, pour pouvoir comparer équitablement les quatre critères, seuls les essais comportant des mesures de reliquat récolte ont été pris en compte.

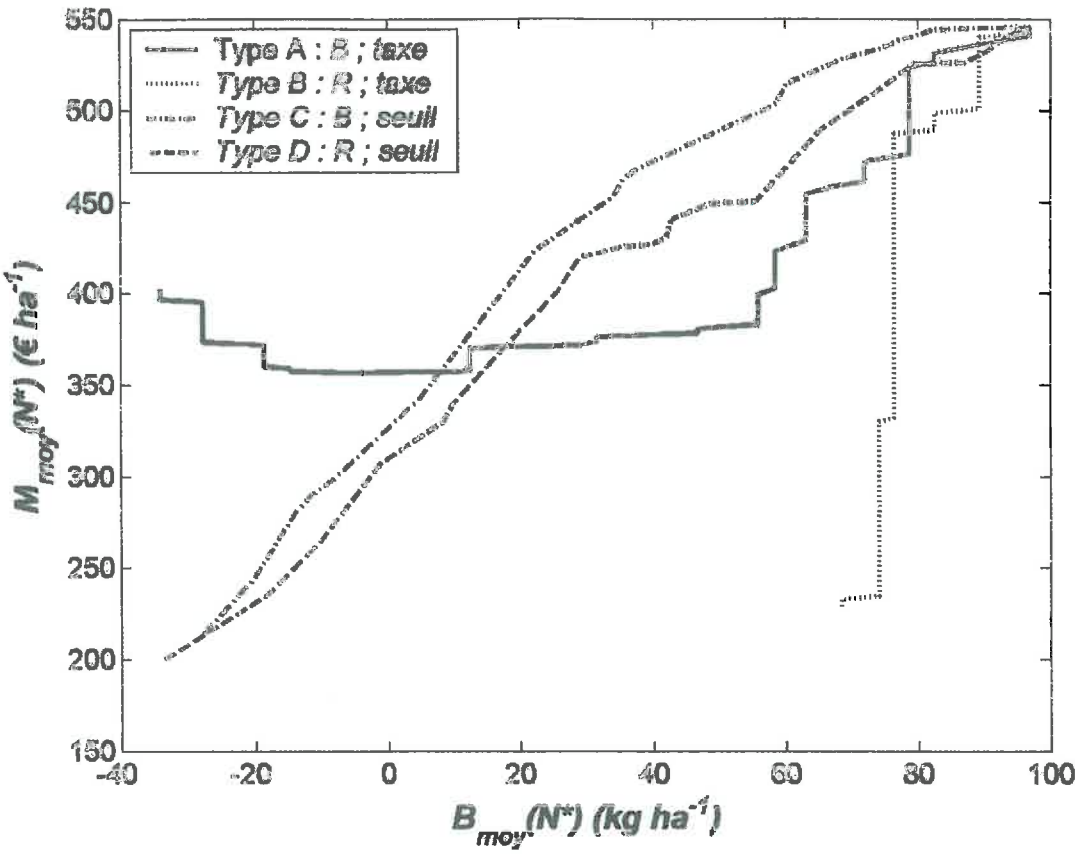


Figure II.5-7 : Confrontation des valeurs de bilan d'azote et de marge brute obtenues aux meilleures doses pour les différents critères et pour une évolution continue de α, β, γ et δ. Seuls les essais comportant des mesures de reliquats à la récolte ont été utilisés.

On retrouve le fait que le critère B ne permet pas d'atteindre des valeurs de bilan d'azote inférieures à 60 kg ha^{-1}. Des trois critères restants, le critère A est celui qui, pour une même valeur de bilan d'azote atteint, donne toujours la marge brute la plus faible jusqu'à des valeurs très basses de bilan d'azote (pour des contraintes environnementales très fortes). Le critère C est le plus intéressant jusqu'à ces faibles valeurs de bilan d'azote. Le critère D occupe une position intermédiaire. Ce n'est que lorsque le bilan d'azote devient négatif et que la taxe du critère A devient positive, s'apparentant ainsi à une subvention, qu'il devient le plus intéressant. Cette analyse confirme que (i) les meilleurs critères combinant des impératifs économiques et environnementaux sont basés sur des seuils et non des taxes et (ii) que les marges obtenues avec une contrainte environnementale sont très largement inférieures à celles obtenues sans contrainte, et que la préservation de l'environnement a un coût important.

II.5.4 Conclusion partielle

Les critères à seuil sont préférables pour combiner au mieux intérêts environnementaux et économiques. Néanmoins, il conviendrait d'envisager d'éventuelles sanctions à appliquer en cas de dépassement du seuil fixé. Pour ce qui est du choix de la variable environnementale prise en compte, bilan d'azote ou reliquats, il se porte sur la première pour plusieurs raisons : (i) le critère C semble donner de meilleurs résultats sur notre jeu de données que le critère D du point de vue environnemental et économique (de même le critère A donne de meilleurs résultats que son homologue B) ; (ii) le reliquat varie assez peu en fonction de la dose d'azote ; (iii) le bilan d'azote est mieux simulé que les reliquats à la récolte par STICS.

Dans la suite, seuls les critères A et C, basés sur la variable environnementale bilan d'azote, seront considérés. Le critère basé sur une taxe est conservé dans la suite de l'analyse eu égard au projet de loi qui envisageait d'employer un tel critère.

II.6 Evaluation de l'aptitude de STICS à déterminer une stratégie de fertilisation azotée selon un critère agro-environnemental

Après avoir (i) paramétré et validé le modèle STICS sur les variables d'intérêt – le rendement (G), la teneur en protéines (P), les reliquats à la récolte (R) et le bilan d'azote (B) – (ii) choisi une forme de critère agro-environnemental, nous évaluons dans cette partie la capacité de STICS à préconiser des doses d'azote optimales selon le critère choisi. Les performances des recommandations de STICS seront comparées à celles effectuées par une méthode de référence en la matière, AZOBIL (Machet *et al.*, 1990).

II.6.1 Méthode d'évaluation utilisée

II.6.1.1 Choix d'une méthode

La méthode d'évaluation de la qualité des préconisations de STICS utilisée ici suivra celle décrite dans la partie II.2.3.

En adaptant la démarche proposée par Wallach (2002) à notre situation, on peut proposer trois méthodes de détermination des doses optimales d'après les jeux de données observées, N^* et de données simulées, $\hat{N}^*$. Ces trois méthodes sont illustrées par la Figure II.6-1.

La première méthode ne considère que les doses réellement expérimentées : ainsi, plutôt que d'évaluer la capacité du modèle à préconiser une dose d'azote optimale, c'est en réalité sa capacité à effectuer des *classements de scénarios prédéfinis de fertilisation azotée et à sélectionner le meilleur* qui est jaugée. Cette méthode biaise les résultats dans le sens où le modèle a de fait un choix limité de doses d'azote : il ne peut donc pas préconiser « sa » dose optimale et lorsqu'il se trompe, l'erreur sur la dose est au moins égale au pas entre deux doses expérimentées.

La deuxième méthode consiste à ajuster des modèles de réponse à la dose sur le critère (ici la marge). Ces modèles de réponse permettent de déterminer des doses optimales qui se calculent en fonction des paramètres de la fonction ajustée dans le cas observé et simulé. Cette méthode présente l'avantage d'estimer la dose optimale de façon continue et non discrète comme dans le cas précédent. Mais, si le modèle surestime la réponse à la dose d'azote, la dose préconisée $\hat{N}^*$ risque d'être largement surestimée, comme dans l'exemple

montré. En outre, les valeurs de N^* et de $\hat{N}^*$ seront très dépendantes du modèle d'ajustement choisi (Makowski *et al.*, 2001). Cette façon de faire n'a pas été retenue pour ces raisons.

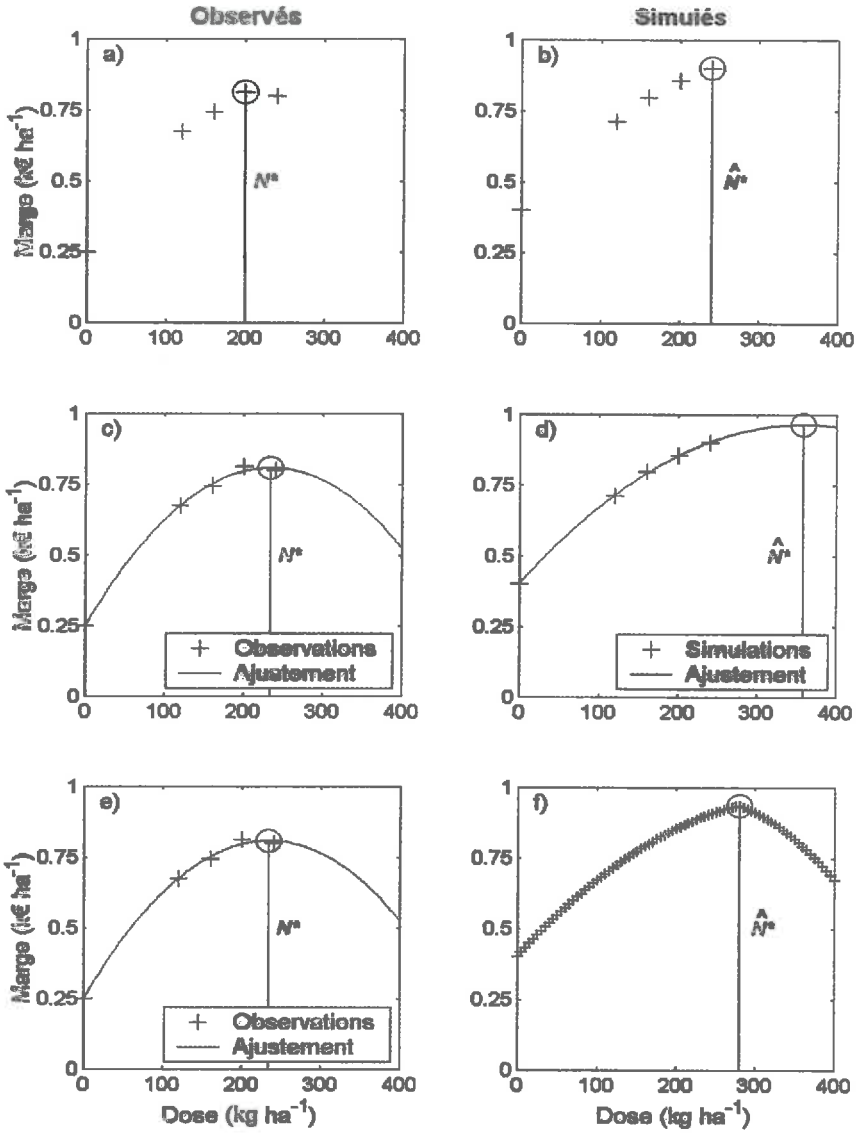


Figure II.6-1 : Illustration de trois méthodes de comparaison de N^* et de $\hat{N}^*$; a) et b) en ne considérant que les doses réellement expérimentées ; c) et d) par ajustement d'une fonction de réponse identique sur les données observées et simulées pour en déduire les doses optimales ; e) et f) même situation que dans le cas précédent pour les données observées mais pour le modèle réalisation d'un grand nombre de simulations avec un pas de doses très faible.

La troisième méthode détermine N^* exactement comme dans le cas précédent. En revanche, $\hat{N}^*$ est calculée en réalisant des simulations à un pas de dose très faible (de l'ordre de 10 kg ha^{-1}), ce qui permet de réellement obtenir la dose optimale du modèle. Mais dans

cette dernière méthode, on compare de ce fait des doses optimales observées et simulées qui sont déterminées de façon très dissemblable. Nous avons donc préféré utiliser la première.

On pourra par la suite parler de « dose optimale », mais il faudra entendre « la meilleure dose parmi les doses testées ». Elles seront également notées N^* et $\hat{N}^*$ respectivement pour la meilleure dose réelle parmi les doses expérimentées et la meilleure dose selon le modèle.

Pour estimer la capacité du modèle à préconiser des doses d'azote, nous comparerons donc les valeurs moyennes sur les essais :

- de $\overline{N}^*$ et $\overline{\hat{N}}^*$;
- des valeurs de bilans d'azote réellement obtenus avec N^* et $\hat{N}^*$, c'est-à-dire $\overline{B}(N^*)$ et $\overline{B}(\hat{N}^*)$;
- des valeurs de marges réellement obtenues avec N^* et $\hat{N}^*$, c'est-à-dire $\overline{M}(N^*)$ et $\overline{M}(\hat{N}^*)$;
- dans le cas de l'utilisation du seuil (critère C), la fréquence à laquelle le seuil est dépassé.

Afin de réellement tester la capacité du modèle à réaliser des préconisations de doses d'azote avec un caractère prédictif, les simulations utilisées ne seront pas celles qui ont été présentées jusqu'à présent mais des simulations obtenues avec des séries climatiques comme décrit ci-dessous.

II.6.1.2 Simulations avec des séries climatiques

Des séries climatiques ont été utilisées à partir de la date du deuxième apport, considérant que c'est à ce moment que la préconisation doit être réalisée (dans la seconde moitié du mois de mars). Les séries sont de 27 années à Boigneville, 28 à Thibie et à La Cheppe (poste de Fagnières), 30 à La Minière (poste de Versailles) et 30 à Chambry (poste de Saint-Quentin). Ces données sont issues de stations de Météo France ou de l'INRA.

Pour chacune des variables utilisées pour calculer le critère, la moyenne des simulations effectuées avec les différentes séries climatiques a été effectuée. Pour certaines

variables, comme le bilan d'azote B, on aurait pu utiliser par exemple le quatrième quintile, dans une perspective de gestion des risques. On pourrait ainsi, dans l'utilisation du critère C, rejeter toutes les doses conduisant à excéder le seuil fixé pour B dans plus de 80% des séries. Mais nous avons retenu la moyenne pour la comparer aux données expérimentales.

II.6.2 Validation des simulations réalisées avec des séries climatiques

La Figure II.6-2 est l'équivalente de la Figure II.4-4 mais avec les simulations issues de séries climatiques (les données du paramétrage, Chambry, ont donc été écartées). On constate que la qualité des simulations n'est pas affectée sur notre jeu de données par la méconnaissance du climat postérieur au deuxième apport. Au contraire, les RMSE de G et P

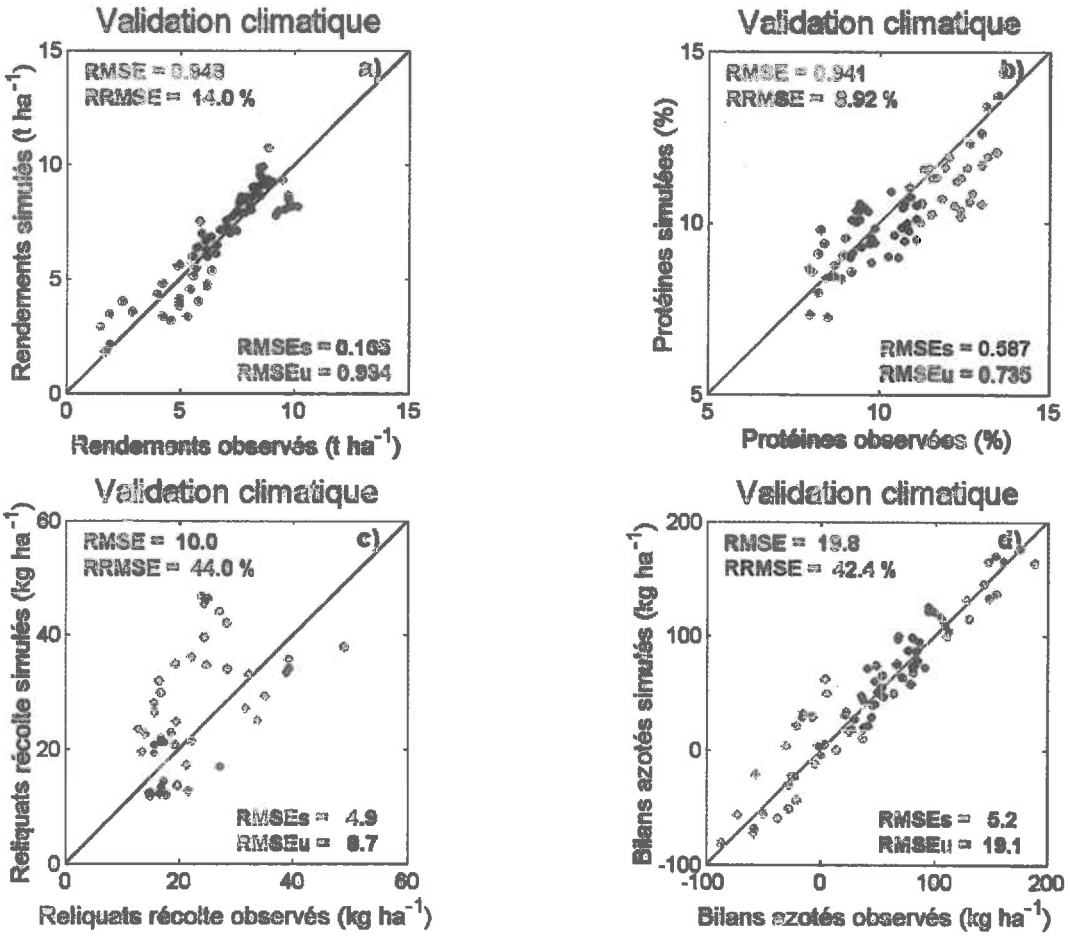


Figure II.6-2 : Valeurs de différentes variables simulées et observées obtenues sur les données ayant servi à réaliser la validation avec utilisation de séries climatiques à partir des dates du deuxième apport : a) Rendement ; b) Teneur en protéines ; c) Reliquat d'azote à la récolte ; d) Bilan d'azote.

auraient tendance à être améliorées : celle du rendement passe de 1.06 à 0.95 t ha^{-1}, celle des protéines, de 1.60 à 0.94%. L'erreur commise sur les variables environnementales ne varie pas : la RMSE du bilan d'azote reste à 10-11 kg ha^{-1} et celle du reliquat reste à environ 20 kg ha^{-1}. La faible différence entre simulations issues des données climatiques observées et issues de séries climatiques peut être attribuée à une assez faible variabilité du climat au cours du printemps dans les zones considérées, et au fait que les années utilisées sont assez proches d'une année climatique « moyenne ».

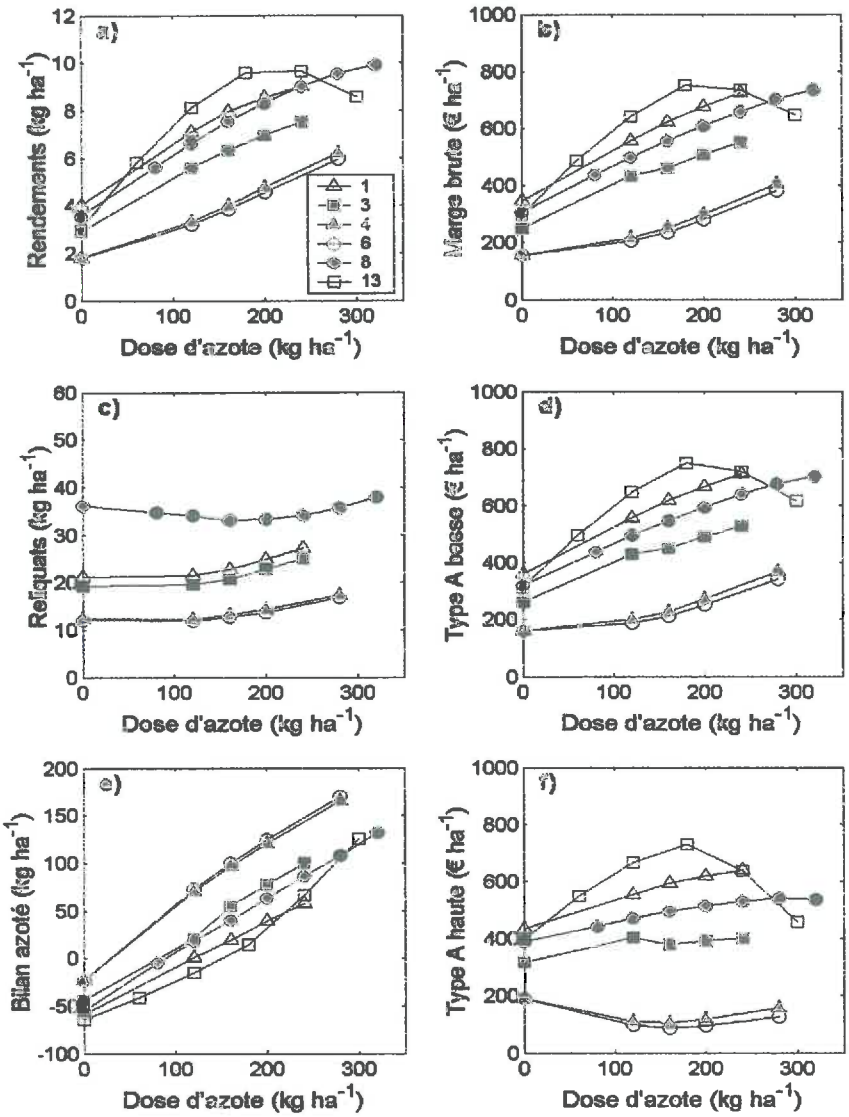


Figure II.6-3 : Courbes de réponse simulées à partir des séries climatiques sur 6 essais : a) Rendement ; b) Marge brute ; c) Reliquat d'azote ; d) Critère A avec α=0.23 € kg^{-1} ; e) Bilan d'azote ; f) Critère A avec α=1.50 € kg^{-1}. Les numéros des essais renvoient au Tableau II.3-6.

Une autre façon de voir les résultats est de considérer sur quelques exemples l'évolution des différentes variables d'intérêt en fonction de la dose d'azote. La Figure II.6-3 est l'équivalente de la Figure II.5-1 mais avec les données issues des simulations obtenues avec les séries climatiques au lieu des données réelles.

Cette figure permet de constater que la réponse du rendement à la dose d'azote est très souvent surestimée : alors que la Figure II.5-1 sur les données observées faisait clairement apparaître un plateau sur les essais 1, 3, 4, 6, celle-ci n'en fait pas apparaître sur les données simulées. L'essai 8 qui tendait à atteindre un plateau pour les plus fortes doses sur les rendements observés n'en présente pas du tout sur les données simulées ; quant à l'essai numéro 13, il fait figure d'exception : alors qu'un plateau assez peu marqué se dessinait sur les données observées, les données simulées ont une forme en cloche. Ceci est lié à ce que les propriétés de sol estimées sur la placette ayant reçu la plus forte dose sont différentes et moins favorables que celles de la placette ayant reçu la dose immédiatement inférieure. La tendance est inverse pour le reliquat : leur réponse était, conformément aux observations classiques, assez plate pour les doses les plus faibles avant de croître fortement. Avec les données simulées, on n'observe pas cette augmentation finale, ce qui est dû à l'absence de plateau atteint sur les rendements. La réponse du bilan d'azote est nettement mieux simulée, sachant que c'est en partie lié à la présence de la dose d'azote dans le calcul de cette variable. Les remarques qui ont été faites sur le rendement sont directement transposables sur les marges et critère agro-environnementaux représentés : les données observées font apparaître un plateau ou une diminution selon les cas ; dans les données simulées, ce plateau ou cette chute n'apparaît pas ou plus tard.

Ces considérations auront un poids prépondérant sur les développements ultérieurs et en ont rétroactivement sur le choix de la façon d'établir les doses optimales observées, N^*, et simulées, $\hat{N}^*$ (voir II.6.1.1). Si en effet le modèle ne fait pas apparaître de plateau, la dose optimale déterminée sur les données simulées par ajustement risque d'être largement surestimée. Ainsi, ne considérer que les doses réellement expérimentées constitue une sorte de garde-fou, $\hat{N}^*$ ne pouvant pas excéder la dose maximale expérimentée.

II.6.3 Evaluation de la méthode pour deux niveaux de taxation ou de seuil

Dans cette partie, nous reprenons les critères A et C avec les valeurs de paramètres adoptées dans le Tableau II.5-2. Nous comparerons les préconisations basées sur les

simulations prédictives du modèle et réalisées selon ces critères aux préconisations d'une méthode de référence en France en terme de fertilisation azotée, AZOBIL (Machet *et al.*, 1990). Le Tableau II.6-1 donne les résultats de l'évaluation de la capacité du modèle à réaliser des préconisations de fertilisation azotée et les compare à ceux obtenus par les doses bilan préconisées par AZOBIL (pour les doses préconisées par ce modèle, *cf.* Tableau II.3-6).

	Critère	STICS	AZOBIL
$\overline{N}^* - \overline{\hat{N}}^*$ (kg ha^{-1})	Type A-b	-32	25
	Type A-h	10	-23
	Type C-b	1	-13
	Type C-h	8	-63
$\overline{B}(N^*) - \overline{B}(\hat{N}^*)$ (kg ha^{-1})	Type A-b	-24	13
	Type A-h	1	-17
	Type C-b	-4	-9
	Type C-h	4	-37
$\overline{M}(N^*) - \overline{M}(\hat{N}^*)$ (€ ha^{-1})	Type A-b	12	14
	Type A-h	50	15
	Type C-b	31	-8
	Type C-h	30	-59
Fréquence de dépassement du seuil	Type C-b	14%	29%
	Type C-h	14%	79%

Tableau II.6-1 : Différences entre (i) les moyennes des meilleures doses réelles et préconisées ; (ii) les bilans azotés moyens réels obtenus avec les meilleures doses réelles et préconisées ; (iii) les marges moyennes réelles obtenues avec les meilleures doses réelles et préconisées ; (iv) la fréquence avec laquelle la dose préconisée conduit à excéder en réalité le seuil fixé ; pour les critères A et C, hypothèses basse et haute. Les résultats sont indiqués pour les doses préconisées par STICS et AZOBIL.

Les doses recommandées par STICS sont en moyenne supérieures aux doses optimales réelles pour le critère A dans l'hypothèse basse. Elles sont en revanche sous-estimées dans les autres cas. C'est le contraire pour AZOBIL : les critères agro-environnementaux assez contraignants imposent des doses plus faibles que celles sélectionnées par ce modèle adapté à des contraintes faibles, sauf dans le cas de faible contrainte environnementale : AZOBIL semble être plus environnementaliste que le critère A en hypothèse basse.

Les bilans obtenus avec les doses préconisées par STICS sont proches de ceux obtenus avec les vraies doses optimales dans la plupart des cas, sauf le critère A, hypothèse basse, ce qui est cohérent avec ce qu'on a vu sur les doses. AZOBIL présente cette fois-ci encore un comportement inverse : alors qu'il présente un bilan meilleur que ce qu'on aurait dû obtenir avec le type A-b, dans toutes les autres configurations, il présente un moins bon bilan environnemental, ce qui est lié au fait qu'il n'est pas prévu pour de telles contraintes.

Au niveau des marges, dans le cas du critère A, la différence calculée ne peut par construction qu'être positive ou nulle, c'est-à-dire qu'on ne peut avoir que des pertes. Pour le critère C, cette différence peut être négative, si les doses retenues sont trop importantes, générant des revenus supérieurs mais se traduisant par un bilan d'azote qui excède le seuil objectif. Pour le type A-b, proche de la configuration actuelle, STICS et AZOBIL conduisent à des pertes équivalentes. Ce qui fait que pour ce critère, AZOBIL peut être considéré meilleur que STICS puisque par ailleurs son bilan d'azote est plus favorable. Pour les autres critères, les pertes générées par STICS peuvent être très importantes. Dans le critère A-h, il atteint l'objectif environnemental ($\overline{B}$ (N*) - $\overline{B}$ ($\hat{N}$*) proche de 0), mais ceci se traduit par une perte moyenne de 50 € ha^{-1}. Pour C-b et C-h, AZOBIL dégage des marges supérieures aux marges obtenues avec la dose optimale vraie. Ceci est lié à ce que les objectifs environnementaux ne sont pas atteints, comme en témoigne le fort taux de dépassement du seuil imposé au bilan d'azote. Ici encore, STICS atteint mieux les objectifs environnementaux qu'AZOBIL, mais au prix de pertes de revenus importantes.

II.6.4 Evaluation de la méthode pour une évolution continue des valeurs de taxation et de seuil

La Figure II.6-4 présente l'évolution des valeurs de $\overline{N}^* - \overline{\hat{N}}^*$, de $\overline{G}$ (N*) - $\overline{G}$ ($\hat{N}$*), de $\overline{B}$ (N*)- $\overline{B}$ ($\hat{N}$*) et de $\overline{M}$ (N*) - $\overline{M}$ ($\hat{N}$*) en fonction du paramètre α pour le critère A pour STICS et AZOBIL. La Figure II.6-5 fait de même pour le paramètre γ et le critère C.

En ce qui concerne le critère A et l'évaluation de STICS, trois phases peuvent être dégagées. Au cours de la première, pour des valeurs de α allant de 0 à 1.5 environ, le modèle sur-estime les doses préconisées par rapport à la vraie meilleure dose. Ceci a peu de conséquences sur le rendement puisque la différence $\overline{G}$ (N*) - $\overline{G}$ ($\hat{N}$*) est sur cet intervalle très proche de 0. Mais cela se traduit en revanche par un moins bon bilan environnemental qu'à l'optimum vrai. Une phase de transition suit, pour α allant de 1.5 à 3 environ. Les doses

choisies par STICS correspondent alors à peu près aux meilleures doses, sauf à la fin de cet intervalle. Le bilan environnemental correspond alors au bilan atteint avec les doses correctes au début de cet intervalle. Au cours de la troisième phase, les doses préconisées par STICS parmi celles proposées sont systématiquement sous-estimées. Ceci se traduit par un bilan environnemental qui est meilleur que le bilan obtenu avec les vraies doses optimales, mais par une perte de rendement. Au niveau des pertes de revenus, c'est au cours de la deuxième phase qu'ils sont les plus importants, plafonnant aux alentours de 50 € ha^{-1}. En ce qui concerne AZOBIL, on voit très clairement que pour les très faibles contraintes environnementales, il a tendance à préconiser des doses pas assez fortes, ce qui se traduit par de meilleurs bilans et par une très légère perte de rendement et de marge. Après, au fur et à mesure que la contrainte augmente, AZOBIL préconise des doses trop fortes par rapport aux doses optimales réelles, ce qui se traduit par un gain de rendement, une dégradation du bilan environnemental par rapport aux doses optimales et une très forte perte de marge liée au poids de la taxe sur le bilan

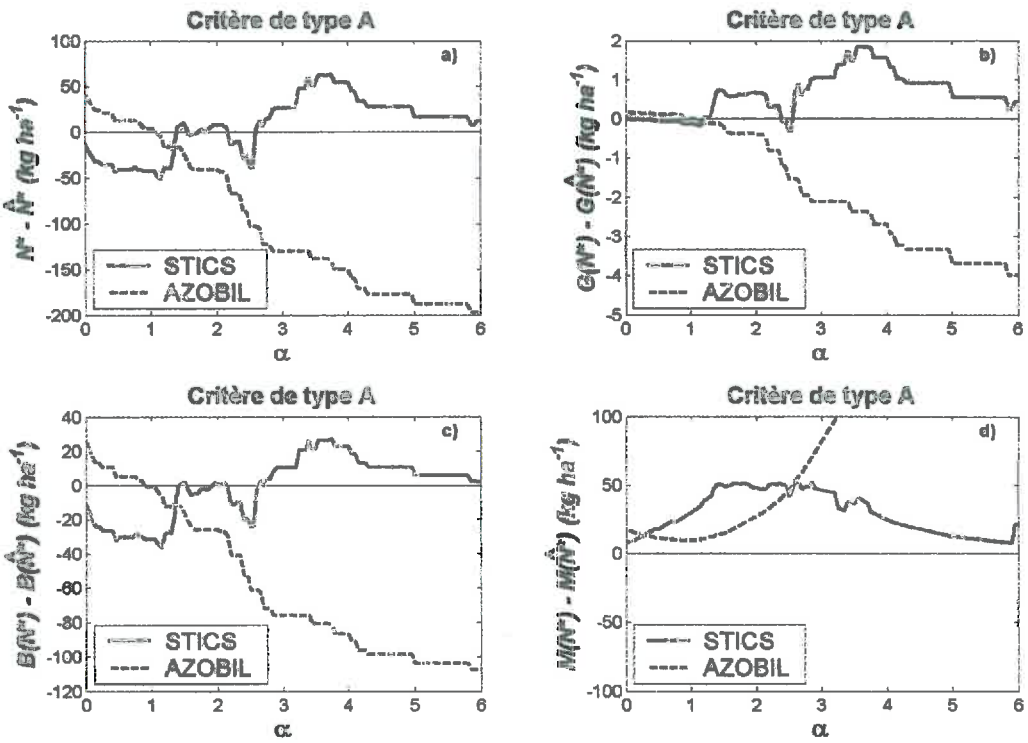


Figure II.6-4 : Comparaison des performances respectives des préconisations de STICS et d'AZOBIL par rapport au critère A pour une gamme de valeurs de α allant de 0 à 6 ; a) Différence des moyennes sur les essais des vraies meilleures doses (N^*) et des doses préconisées ($\hat{N}^*$) ; b) Différence des moyennes des rendements obtenus avec les vraies meilleures doses (N^*) et les doses préconisées ($\hat{N}^*$) ; c) Différence des moyennes des bilans azotés obtenus avec les vraies meilleures doses et les doses préconisées ; d) Différence des moyennes des marges brutes obtenues avec les vraies meilleures doses et les doses préconisées.

d'azote. AZOBIL, d'après ce jeu de données, semble être adapté à préconiser des doses optimales selon le critère A pour une valeur de α=1.

La Figure II.6-5 présente l'évolution de ces mêmes différences pour le critère C. Ici aussi, les axes des abscisses sont inversés, de façon à ce que la contrainte environnementale augmente quand on se déplace vers la droite. On peut d'abord remarquer que pour STICS, lorsque γ varie, les quatre différences représentées sont toujours plus proches de 0 que ce n'était le cas avec le critère A. Par exemple, le maximum d'écart de dose est de 25 kg ha^{-1} et le minimum de –23 pour le critère C alors qu'il était de 63 pour le maximum et de –51 pour le minimum pour le critère A. Autrement dit, STICS parvient plus à se rapprocher des meilleurs choix effectués *a posteriori* lorsque ceux-ci sont fait avec le critère C qu'avec le A. Une valeur caractéristique de γ pourrait être 75, au-dessus de laquelle STICS surestime assez régulièrement les préconisations de dose, ce qui se traduit par une dégradation du bilan d'azote par rapport à l'optimum et à une légère dégradation de la marge, constante. Puis, entre 40 et 75 environ, il sous-estime les doses à apporter, ce qui se traduit par un meilleur bilan d'azote qu'à l'optimum mais par une perte de rendement et financière. En-dessous de 40, soit pour les valeurs les plus contraignantes, STICS se remet à préconiser des doses trop fortes, occasionnant des rendements et des marges plus forts que ceux qui auraient dû être obtenus. Mais, de façon corollaire, les objectifs par rapport au bilan environnemental ne sont pas atteints. AZOBIL quant à lui sous-estime les doses à apporter jusqu'à γ=100, contrainte à partir de laquelle les doses qu'il préconise ne satisfont pas au critère environnemental. Ainsi, de même que les préconisations d'AZOBIL semblaient d'après notre jeu de données être adaptées à une valeur de α=1 pour le critère A, elles semblent être adaptées à une valeur de γ=100 pour le critère C. On peut d'ailleurs remarquer que pour cette valeur de γ, STICS préconise également les bonnes doses.

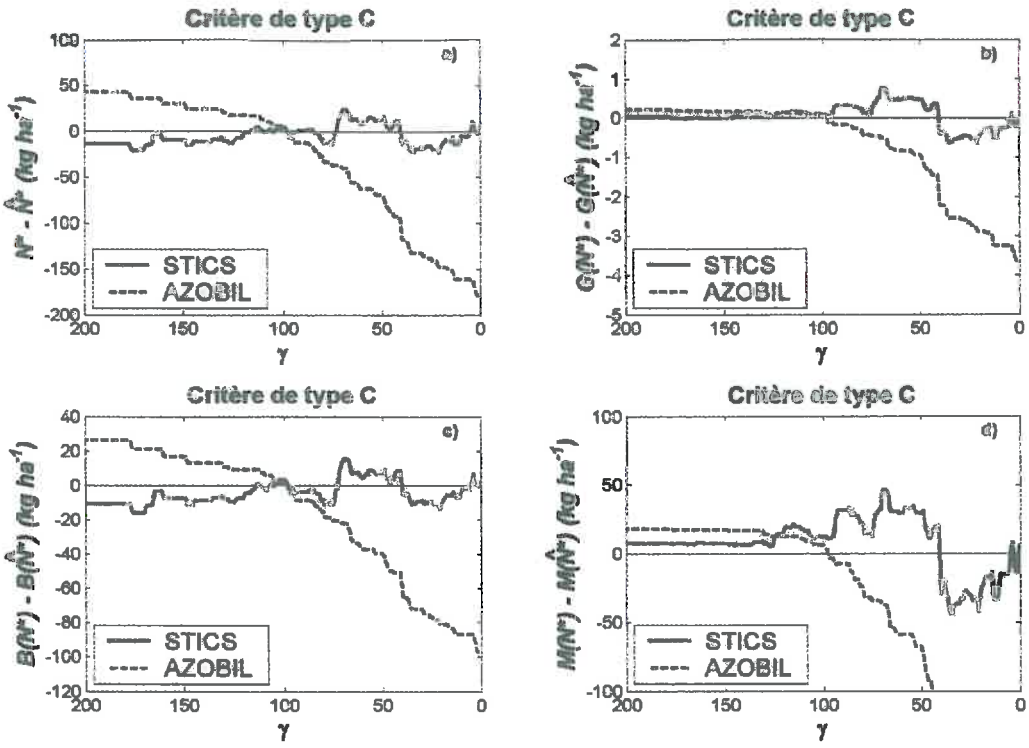


Figure II.6-5 : Comparaison des performances respectives des préconisations de STICS et d'AZOBIL par rapport au critère C pour une gamme de valeurs de γ allant de 0 à 200 ; les axes des X sont inversés de façon à ce que la contrainte environnementale augmente en allant vers la droite ; a) Différence des moyennes sur les essais des vraies meilleures doses et des doses préconisées ; b) Différence des moyennes des rendements obtenus avec les vraies meilleures doses (N^*) et les doses préconisées ($\hat{N}^*$) ; c) Différence des moyennes des bilans azotés obtenus avec les vraies meilleures doses et les doses préconisées ; d) Différence des moyennes des marges brutes obtenues avec les vraies meilleures doses et les doses préconisées.

II.7 Conclusion

Parmi les quatre critères proposés, le critère C semble se révéler le plus intéressant pour effectuer des préconisations de fertilisation azotée à l'aide de STICS pour plusieurs raisons :

- on obtient de meilleurs résultats du point de vue agro-environnemental comme il a été montré sur les données expérimentales. D'une part, l'utilisation de la variable bilan d'azote semble plus pertinente que celle de la variable reliquats à la récolte car ceux-ci varient peu en fonction de la dose d'azote et n'incluent pas les pertes gazeuses ; d'autre part, l'utilisation de seuil permet de réduire l'impact environnemental avec des conséquences plus faibles sur la marge de l'agriculteur que les critères à taxe. Il faut cependant remarquer qu'on n'a pas inclus de pénalité dans le calcul de la marge des critères à seuil dans les cas où le seuil est dépassé ;

- le bilan d'azote est mieux simulé par le modèle que le reliquat à la récolte ;

- le modèle réalise de façon générale de meilleures préconisations lorsqu'on utilise le critère C que lorsqu'on utilise le A pour une large gamme des paramètres α et γ.

Le critère C sera donc utilisé dans la partie consacré au choix de doses à l'échelle de la parcelle. L'utilisation d'un tel critère présente l'intérêt qu'elle équivaut dans une large mesure à utiliser la capacité du modèle à classer la valeur simulée du bilan d'azote par rapport à un seuil, ce qui est plus facile que de l'évaluer précisément.

Par rapport à la préconisation de doses, le modèle a un moins bon comportement qu'AZOBIL pour les faibles contraintes environnementales mais, globalement, la qualité de ses préconisations semblent à peu près satisfaisante pour une large gamme de contraintes. Afin de se faire une meilleure idée de cette qualité, il faudrait appliquer la même analyse avec d'autres modèles ou d'autres méthodes de choix. D'autres méthodes d'évaluation de la qualité de préconisations pourraient également être utilisées, mais elles ne sont pas légions.

Les principaux résultats de cette partie sont repris dans Houlès *et al.* (accepté).

III - Troisième Chapitre
Elaboration de deux approches de spatialisation du modèle de culture à l'échelle intra-parcellaire

Que les hommes ne soient donc pas trop assurés,
quand ils jugent, comme celui qui croit
que les blés sont mûrs avant qu'ils le soient

Dante, La Divine Comédie, Le Paradis, XIII, 130-132

III.1 Introduction 129
III.2 Spatialisation du modèle basée sur un inventaire cartographique des sols 130
III.2.1 Introduction 130
III.2.2 Cartographie des sols et liens avec un modèle de culture 131
III.2.3 Présentation des données utilisées 142
III.2.4 Caractérisation spatiale des entrées sol de STICS 146
III.2.5 Conclusion 156
III.3 Spatialisation du modèle par assimilation de données 157
III.3.1 Objectif et principe de l'assimilation de données issues de la télédétection dans le cadre de l'agriculture de précision 157
III.3.2 Présentation des trois méthodes d'assimilation utilisées 160
III.3.3 Comparaison des trois méthodes sur un jeu de données simulées 178
III.3.4 Analyse de différentes configurations d'assimilation avec la méthode GLUE 205
III.3.5 Discussion et conclusions 217
III.4 Conclusion 221

III.1 Introduction

Le deuxième volet de la construction d'un outil de préconisation de la fertilisation azotée à l'échelle intra-parcellaire concerne les méthodes de spatialisation du modèle de culture utilisé. Ceci permettra de réaliser aux dates d'apport un diagnostic spatialisé de l'état de nutrition azotée de la culture et, en appliquant les méthodes vues dans le chapitre précédent, on sera en mesure de prédire des doses d'azote optimales en tout point de la parcelle.

Nous avons vu dans le chapitre introductif qu'il existe au moins deux grandes approches permettant de réaliser des simulations spatialisées. La première s'appuie sur l'utilisation de bases de données cartographiques et pédologiques. La deuxième consiste à utiliser des variables dérivées des mesures de télédétection, elles-mêmes spatialisées, comme entrées d'un système d'assimilation d'informations dans le modèle.

Une revue bibliographique des méthodes permettant d'établir des cartes pédologiques sera réalisée avant de décrire la construction d'une couverture pédologique associée à des règles de pédotransfert sur nos deux parcelles expérimentales. Couverture pédologique et règles de pédotransfert nous permettront de renseigner le modèle en tout point de la parcelle.

Trois méthodes d'assimilation de données issues de la télédétection seront ensuite présentées et comparées sur des jeux de données virtuels, de façon à en retenir une qui sera par la suite utilisée sur les jeux de données réels. Quelques caractéristiques requises concernant les données de télédétection seront également dégagées à cette occasion.

A l'issue de ce chapitre méthodologique, deux méthodes de spatialisation de STICS à l'échelle intra-parcellaire seront en place : la première par approche cartographique, la deuxième par assimilation de données issues de la télédétection. L'objet du chapitre suivant sera de les évaluer et de les comparer sur le jeu de données expérimental et de mettre en œuvre les méthodes de préconisation spatialisée de doses d'azote.

III.2 Spatialisation du modèle basée sur un inventaire cartographique des sols

L'objectif de cette partie est d'appliquer aux parcelles expérimentales de Chambry une méthode permettant de constituer des cartes à l'échelle intra-parcellaire de chacune des variables d'entrée sol du modèle basée sur une prospection pédologique. Une telle approche est très coûteuse en mesures et n'est guère envisageable dans un cadre opérationnel. Elle servira ici de référence à la méthode utilisant l'assimilation des données issues de la télédétection. Elle comprend deux étapes : (i) la caractérisation spatiale des propriétés du sol par un inventaire pédologique et (ii) l'établissement de relations entre la description des caractéristiques du sol ainsi obtenues et les variables d'entrées du modèle.

III.2.1 Introduction

Une analyse bibliographique des méthodes concernant ces deux aspects de la constitution de couvertures pédologiques permettra de dégager un certain nombre de règles spécifiques de la cartographie, à des échelles compatibles avec les nécessités de l'agriculture de précision, notamment en terme de résolution spatiale. Puis, le modèle STICS ayant été décrit dans le Chapitre II, nous nous bornerons à rappeler la nature des informations sur le sol nécessaires aux simulations. L'ensemble des mesures réalisées sera présenté, ainsi que les méthodes permettant une caractérisation spatiale de la variabilité des sols. La description de la constitution des cartes pédologiques des parcelles expérimentales ne sera que succinctement esquissée, étant donné qu'elle n'a pas été réalisée dans le cadre direct de notre étude, mais dans celui du projet « Agriculture de Précision » de l'INRA, en collaboration avec l'unité d'Orléans. L'établissement de règles de pédotransfert locales, interfaces entre les cartes et le modèle STICS, sera décrit plus exhaustivement et, enfin, un point sera fait sur les variables dont la caractérisation spatiale a nécessité un krigeage.

III.2.2 Cartographie des sols et liens avec un modèle de culture

Les deux étapes de cette approche correspondent à la cartographie des sols et à l'établissement de fonctions de pédotransfert.

III.2.2.1 Cartographie des sols

III.2.2.1.1 Origine des cartes pédologiques

La cartographie des sols est une discipline ancienne. En France par exemple, elle remonte aux années 1940 (Jamagne *et al.*, 1995). Elle se base sur une classification des sols, une typologie, qui décrit essentiellement l'origine géologique du sol, le processus de pédogenèse qui l'a formé et les caractéristiques des différents horizons : matériau et propriétés. On peut citer les classifications américaine de 1960, française de 1967 et de la FAO de 1973 (Duchaufour, 1977). Classiquement, une carte pédologique se compose des éléments suivants : (i) le dessin, composé du fond topographique, des contours et des graphismes indiquant la nature du ou des sols à l'intérieur de chaque unité ; (ii) la légende

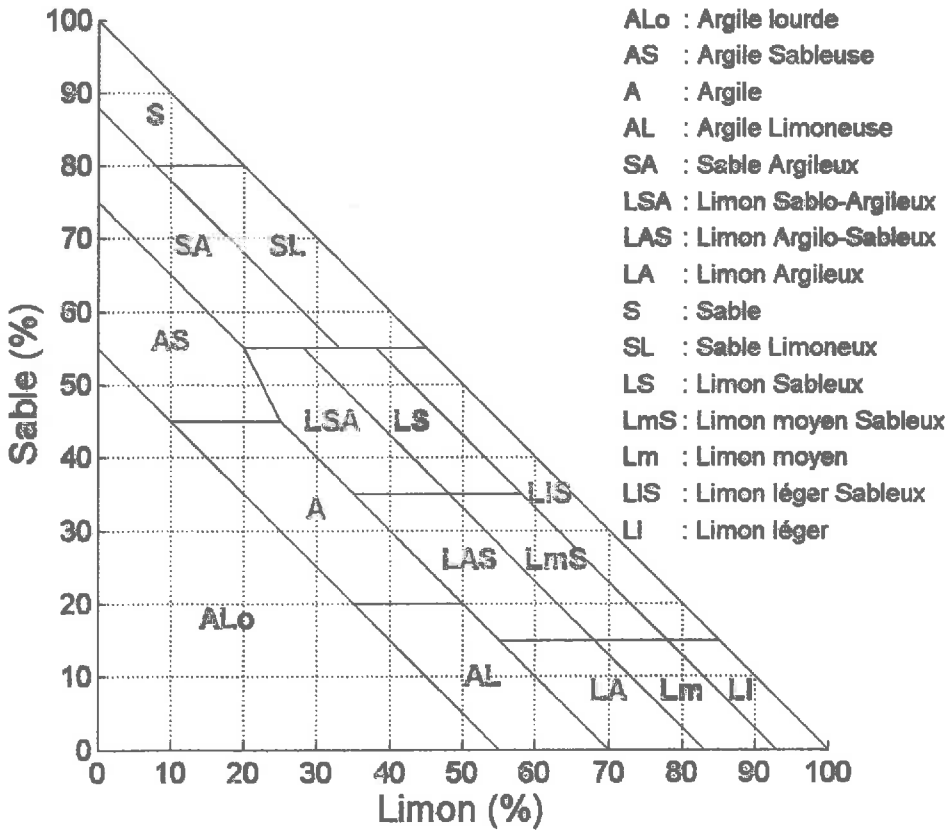


Figure III.2-1 : Triangle des textures de la carte des sols de l'Aisne (Jamagne, 1967). Les classes de textures sont celles qui sont utilisées dans notre étude.

établie d'après une classification de référence des différentes unités cartographiques ; (iii) la notice ou le rapport qui a pour rôle d'approfondir la légende (Boulaine, 1980). La légende fait référence à une classification des matériaux en fonction de leur granulométrie donnée par classes et qui peut être représentée par un triangle des textures. Nous reproduisons ici celui qui est associé à la carte des sols de l'Aisne (Jamagne, 1967) et qui servira de référence dans notre étude. La légende peut également contenir des informations sur la teneur en calcaire du sol ou la richesse en matière organique du premier horizon.

L'objectif initial de ces cartes était de mettre en évidence la répartition et les caractéristiques des sols et d'interpréter ces propriétés en terme de comportement, notamment hydrique. Le but pouvait déjà être d'établir des diagnostics comme des risques de sécheresse (Jamagne *et al.*, 1977). Mais ce qui fait leur intérêt par rapport à l'utilisation dans des modèles de culture est qu'elles font appel à des concepts, comme la réserve hydrique des sols, qui sont utilisés dans nombre d'entre eux.

III.2.2.1.2 Informatisation des cartes pédologiques et notion d'unité cartographique

L'informatisation des cartes existantes à toutes les échelles a été réalisée en les intégrant dans des systèmes d'information géographique (SIG) au cours des années 1980 en Europe (Jamagne *et al.*, 1995). Les SIG sont un élément indispensable en agriculture de précision (Pierce *et al.*, 1999) et sont largement utilisés dans les applications actuelles de modèles de culture spatialisés, quelles qu'en soient les échelles (par exemple, Heinemann *et al.*, 2002, Jones *et al.*, 2003). Une carte pédologique sous SIG a la structuration suivante : elle est découpée en polygones qui correspondent à des Unités Cartographiques de Sol (UCS) qui sont elles-mêmes composées de plusieurs Unités Typologiques de Sol (UTS). Pour chaque UCS, le pourcentage de chaque UTS qui la compose est donné. Les UTS correspondent à des profils, volumes constitués de la superposition d'horizons aux caractéristiques bien définies selon une classification de référence (Jamagne *et al.*, 1995, Lagacherie *et al.*, 2000). Les règles de pédotransfert définies ci-après permettront de passer de ces caractéristiques aux variables d'entrée du modèle.

III.2.2.1.3 Echelles des cartes existantes

Une caractéristique importante de ces cartes des sols du point de vue de l'agriculture de précision est l'échelle. On peut distinguer quatre grandes catégories (Boulaine, 1980) : (i) les très petites échelles : au-delà de 1:1 000 000 utilisées pour les cartes réalisées au niveau mondial ; (ii) les petites échelles du 1:1 000 000 au 1:250 000, pour les cartes nationales ; (iii)

les moyennes échelles du 1:200 000 au 1:50 000, qui donnent des cartes détaillées situant avec précision l'ensemble des sols mais avec un certain regroupement. Au niveau français, le 1:100 000 est privilégié pour couvrir l'ensemble du territoire ; (iv) les grandes échelles du 1:25 000 au 1:10 000, beaucoup plus rares, utilisables pour des projets précis ou pour le conseil aux agriculteurs. Si les petites ou moyennes échelles peuvent être utiles pour les approches de modélisation au niveau national ou régional, dans le cadre de l'agriculture de précision, seules les grandes échelles sont susceptibles d'apporter la précision requise. Mais elles ne couvrent actuellement pas encore l'ensemble du territoire, ne concernant que des secteurs de référence de petite superficie. La référence de ce type en France est la carte des sols du département de l'Aisne (Jamagne, 1967) établie au 1:25 000 dans lequel se trouvent justement les parcelles expérimentales.

La résolution des cartes peut dès lors être vue au travers du nombre d'UCS présentes sur une parcelle. A titre d'exemple, la carte des sols de l'Aisne permet de délimiter 4 ou 5 zones sur les parcelles expérimentales utilisées dans notre étude, de 10 ha chacune environ (*cf. infra*). Autrement dit, son utilisation permettrait de délimiter 4 ou 5 zones de sol différentes susceptibles d'être fertilisées différemment. Cette partition de la parcelle peut d'un point de vue opérationnel être pertinente, notamment par rapport aux capacités des outils de modulation actuels. Mais, selon certains auteurs, cette échelle n'est pas suffisante. Sadler *et al.* (2000) rapportent que pour une parcelle de 6 ha, une échelle de 1:24 000 ne permet pas de bien décrire la variabilité intra-parcellaire du rendement. Cette même échelle permet en revanche de restituer la variabilité des rendements à l'échelle de la région.

III.2.2.1.4 Constitution de cartes à une échelle adaptée à l'agriculture de précision

Face à ce constat, la plupart des auteurs traitant de l'agriculture de précision caractérisent eux mêmes leur parcelle à la résolution idoine par une prospection pédologique, c'est-à-dire par des mesures très localisées dans l'espace. Ceux qui utilisent également des modèles de culture se retrouvent confrontés au choix suivant : (i) réaliser les simulations sur les points de mesure et spatialiser ensuite les sorties ou (ii) spatialiser les données d'entrée du modèle et réaliser les simulations sur des unités spatiales supposées homogènes.

Paz *et al.* (1999) réalisent des simulations sur 224 placettes pour une parcelle de 16 ha et McKinion *et al.* (2001) 201 pour 201 ha. Booltink *et al.* (2001) réalisent de façon similaire leurs simulations sur 65 profils de sol établis sur une parcelle de 6 ha. Les données simulées sont ensuite krigées, c'est-à-dire spatialement interpolées, de façon à obtenir des cartes

continues. Vauclin *et al.* (1983) soulignent la possibilité de kriger des variables sous-échantillonnées en utilisant leur corrélation avec des variables mieux échantillonnées : on parle alors de cokrigeage.

D'autres auteurs préfèrent spatialiser les entrées du modèle et réaliser les simulations sur des entités spatiales considérées comme homogènes, équivalentes aux UCS. Pour délimiter ces zones et regrouper les points de mesures pédologiques, plusieurs façons de procéder se rencontrent. Nijbroek *et al.* (2003) utilisent des zones délimitées à partir de plusieurs sources d'informations : des cartes de rendement, des images aériennes du sol, des images de culture, l'expérience de l'agriculteur et des cartes des sols préexistantes. Vrindts *et al.* (2003) réalisent des analyses de sol selon une grille et les regroupent selon des observations de la culture sous la forme de cartes de rendement et d'images de télédétection. Constatant qu'un échantillonnage intensif n'est pas à la portée des agriculteurs, Francis & Schepers (1997) proposent de sélectionner des zones d'échantillonnage d'après les observations pédologiques antérieures, l'historique de la parcelle, les cartes de rendement et les observations de télédétection, de façon à limiter les analyses et à les rendre plus pertinentes. Engel (1997) propose également de définir des unités homogènes à partir d'observations spatiales de la culture, remarquant que les analyses de sol combinées à la géostatistique sont intéressantes pour des applications scientifiques mais trop chères pour des applications pratiques. Enfin, pour évaluer l'intérêt d'opérer des regroupements entre points de mesure, Basso *et al.* (2001) comparent plusieurs façons de réaliser des simulations. La première consiste à utiliser comme entrée du modèle la moyenne des mesures réalisées sur une grille (une simulation pour une parcelle de 7 ha). La deuxième fait une simulation pour chacun des 52 points de la grille. La troisième consiste à regrouper les simulations en trois classes déterminées par analyse d'images de télédétection de la culture. Les RMSE obtenues pour du soja sont respectivement de 465, 245 et 101 kg ha^{-1}. Ils en concluent qu'il vaut mieux exécuter les simulations sur un petit nombre de zones au sein de la parcelle. Ceci peut être interprété comme le fait qu'il est plus facile de prévoir des moyennes que d'arriver à décrire toute la variabilité, d'où l'intérêt de faire des regroupements.

A ce titre, il est de plus en plus courant de pratiquer de l'échantillonnage dirigé, qui consiste à ne pas faire des mesures sur une grille régulière, mais en des points qui auront été déterminés d'après une connaissance antérieure de la parcelle à caractériser, ce qui revient à utiliser les regroupements décrits ci-dessus. Ceci permet d'améliorer la caractérisation pour un même effort d'échantillonnage (Pocknee *et al.*, 1996 ; Clay *et al.*, 1999 ; McBratney &

Pringle, 1999). Lorsqu'il n'y a pas de connaissance antérieure, on peut même conseiller de réaliser l'échantillonnage selon des transects non alignés et imbriqués (Pettitt & McBratney, 1993).

Il ressort de ce bref aperçu que la constitution des cartes de sols à grande échelle peut impliquer de façon combinée les techniques suivantes :

- mesures pédologiques sur des grilles pouvant être systématiques ou guidées par des observations antérieures ;
- techniques de géostatistiques pour spatialiser les données ;
- utilisation d'observations continues : images de sol, de la culture, cartes de rendement.

De nouvelles méthodes d'observation continue des propriétés des sols se font jour : on peut notamment citer les mesures de résistivité électrique qui permettent d'établir des cartes faisant apparaître des discontinuités et dont les valeurs sont corrélées à des propriétés du sol comme la teneur en eau et la densité apparente (Michot, 2003).

III.2.2.2 Etablissement de fonctions de pédotransfert en agriculture de précision

Cette partie décrit la façon dont on peut faire le lien entre les cartes classiquement utilisées et les entrées sol du modèle par l'utilisation de fonctions ou de règles de pédotransfert. Une abondante bibliographie traite des propriétés de rétention en eau du sol. Les propriétés du sol vis-à-vis de l'enracinement sont beaucoup moins abordées.

III.2.2.2.1 Fonctions, classes et règles de pédotransfert

L'origine de ce concept remonte à 1989 et a été introduite par Bouma (Bastet, 1999). Parmi d'autres auteurs, Pachepsky & Rawls (1999) définissent les fonctions de pédotransfert comme la dépendance entre les propriétés de rétention en eau et de conductivité hydraulique du sol et des paramètres de sols « basiques », accessibles dans les bases de données courantes : composition granulométrique, teneurs en carbone et azote organiques, densité apparente, qui correspondent aux informations contenues dans les cartes des sols. Lin *et al.* (1999) rajoutent à cette définition le fait qu'elles peuvent traduire les caractéristiques simples de sol en des paramètres d'entrées plus complexes utilisés dans des modèles. Bruand *et al.*

(2002) précisent qu'elles constituent des modèles empiriques de la relation entre composition du sol et propriétés de rétention en eau.

On voit d'emblée qu'il s'agit de méthodes permettant de dériver des informations complexes et quantitatives à partir d'observations plus simples, souvent qualitatives.

Bastet (1999) propose une hiérarchie des fonctions de pédotransfert que nous reprenons ici :

- une première approche consiste à estimer à partir des caractéristiques de sol élémentaires les teneurs en eau pour un nombre limité de valeurs de potentiel matriciel (par exemple, à la capacité au champ et au point de flétrissement permanent). Dans ce cas de figure, une fonction de pédotransfert se présente sous la forme suivante de $H_p=f_p(X, \theta_p)$, où H_p désigne la teneur en eau du sol à un potentiel p donné, X les informations caractérisant le sol (texture, matière organique...) et θ_p les paramètres de la fonction f_p caractéristique du potentiel p. Ces paramètres peuvent également dépendre de X, par exemple par classes de textures. Cette approche se retrouve entre autres chez Gupta & Larson (1979), Puckett *et al.* (1985) et Pachepsky & Rawls (1999) ;

- la seconde consiste à évaluer de façon continue la fonction qui associe la teneur en eau du sol au potentiel matriciel. Ici, la fonction de pédotransfert se présente donc sous la forme suivante : $H(p)=g(p,X,\theta)$. La plupart du temps, ce sont les paramètres θ de la fonction mathématique utilisée pour lier la teneur en eau H au potentiel p qui sont liés aux caractéristiques de sol élémentaires X : $\theta=h(X)$. Les fonctions de ce type les plus rencontrées sont celles de Van Genuchten, Brooks & Corey, Campbell, De Jong ... Ce genre de fonctions de pédotransfert pourrait nous être utile mais se heurte à l'inconvénient que la teneur en eau à la capacité au champ, une des entrées de STICS, ne correspond pas à un potentiel donné (Bruand *et al.*, 1996). Elles sont largement utilisées dans les approches les plus récentes qui visent à décrire le fonctionnement hydrique du sol, par exemple, Tietje & Tapkenhinrichs (1993), Kern (1995), Rajkai *et al.* (1996), Gonçalves *et al.* (1997), Scheinost *et al.* (1997), Sinowski *et al.* (1997), Romano & Santini (1997), Schaap *et al.* (1998), Lin *et al.* (1999) et Mayr & Jarvis (1999).

La seconde approche témoigne d'une complexification des fonctions de pédotransfert développées et qui ne se traduit pas, selon Bastet (1999), par une amélioration significative des estimations de teneur en eau.

La première approche, plus adaptée à notre problème vu que STICS ne requiert que les teneurs en eau à la capacité au champ et au point de flétrissement permanent, peut être rapprochée de ce que Jamagne *et al.* (1995) qualifient de « règle de pédotransfert ». D'après leur définition, une règle de pédotransfert permet de passer des données de base à la disposition de l'utilisateur aux paramètres lui permettant de résoudre des problèmes dans un contexte précis sans pour autant utiliser de fonction mathématique.

Bruand *et al.* (2002) apportent une nouvelle distinction en parlant de « classes de pédotransfert ». Cette définition s'applique aux deux types de fonctions de pédotransfert dégagées par Bastet (1999). Elle consiste à évaluer les paramètres θ de ces fonctions non pas de façon continue comme il est classique de le faire, mais de façon discrète, c'est-à-dire en donnant une valeur de θ par classe de sol (définie selon la texture par exemple). Ceci rejoint les concepts de fonctions de pédotransfert « par classe » et « continues » selon qu'elles sont établies avec un groupement préalable ou non introduits par Wösten *et al.* (1990).

Ces notions de classes et de règles de pédotransfert, au demeurant assez proches en pratique, se rapprochent de ce que nous avons dit des notices associées aux cartes des sols. Un exemple fort illustratif est le tableau associé à la carte des sols de l'Aisne (Jamagne *et al.*, 1977) qui donne la correspondance entre réserves en eau et textures : dans cette approche, on a affaire à des classes de pédotransfert caractérisées par des règles de type « *si*... alors ». Le Tableau III.2-1 le reproduit partiellement.

C'est ce type de présentation qui est la plus adaptée à notre problème car elle correspond le mieux aux informations contenues dans une carte des sols. Dans la suite, la plupart des articles cités parlant de fonctions de pédotransfert, nous ferons de même, mais ce qui sera montré pour celles-ci sera applicable aux classes. A ce titre, Wösten *et al.* (2001) et Rawls *et al.* (2001) font également cette distinction puisqu'ils comparent les performances de fonctions de pédotransfert sur des textures d'une part et sur des classes de texture d'autre part. Ceci rejoint ce qui va être dit dans la partie suivante où il est dit qu'il est nécessaire de procéder à des regroupements par classes afin d'améliorer les performances de ces fonctions.

III.2.2.2.2 Méthodologie d'établissement des fonctions de pédotransfert

De très nombreux types de fonctions de pédotransfert ont été proposés, recensés par Bastet (1999). La plupart d'entre eux utilisent comme entrées des informations concernant la texture (classe texturale selon une nomenclature particulière ou granulométrie), la densité apparente et la teneur en matière organique (par exemple, Tietje & Tapkenhinrichs, 1993, Kern, 1995, Rajkai, 1996, Sinowski *et al.*, 1997). Certains rajoutent d'autres informations, comme la porosité, la teneur en cailloux et la profondeur de l'horizon considéré (Schaap *et al.*, 1998). L'abondance de la bibliographie traitant d'ajustement de fonctions de pédotransfert à des jeux de données de sol montre le caractère non universel des valeurs de paramètres obtenus. Bastet (1999) et Gijsman *et al.* (2002) montrent bien la nécessité d'adapter les jeux de paramètres à chaque condition pédologique, voire même la forme des fonctions de

Texture	Humidité (%) à la capacité au champ	Humidité (%) au point de flétrissement	Densité apparente
Sable (S)	8	3	1.35
Sable limoneux (SL)	12	5	1.40
Sable argileux (SA)	19	10	1.50
Limon léger sableux (LlS)	15	7	1.50
Limon sableux (LS)	19	9	1.45
Limon moyen sableux (LmS)	20	9	1.45
Limon sablo-argileux (LSA)	22	11	1.50
Limon argilo-sableux (LAS)	24	12	1.45
Limon léger (Ll)	17	8	1.50
Limon moyen (Lm)	23	10	1.35
Limon argileux (LA)	27	13	1.40
Argile sableuse (AS)	33	22	1.55
Argile (A)	37	25	1.45
Argile limoneuse (AL)	32	19	1.40
Argile lourde sédimentaire	29	18	1.50
Argile lourde altérée	38	25	1.30

Tableau III.2-1 : Table de correspondance entre texture et propriétés de rétention en eau des sols de la Carte des Sols de l'Aisne, d'après Jamagne *et al.*, 1977. Les textures font références au triangle des textures de l'Aisne, présenté dans ce chapitre.

pédotransfert elles-mêmes. Pachepsky & Rawls (1999), en établissant des fonctions de pédotransfert, divisent leur jeu de données de façon à pouvoir appréhender (i) leur « précision », évaluée sur les jeux de données servant à les établir et (ii) leur « fiabilité », évaluée sur les jeux de données de validation. Ils constatent que si la précision obtenue est bonne, la fiabilité ne l'est pas autant, montrant par là la très forte sensibilité de la qualité de l'estimation des propriétés hydriques des sols au jeu de données ayant servi à établir la fonction de pédotransfert utilisée. Wösten *et al.* (2001) concluent quant à eux sur le fait que les fonctions de pédotransfert permettent de réaliser des interpolations et non des extrapolations, autrement dit qu'elles ne sont valables que dans les conditions dans lesquelles elles ont été établies.

Constatant ainsi que les fonctions de pédotransfert existantes ne sont pas adaptées à leurs jeux de données, nombre d'auteurs concluent qu'elles doivent être établies localement. Ceci provient de ce que les caractéristiques généralement prises en compte ne sont pas seules à jouer sur les propriétés de sols que l'on veut déterminer. Bruand *et al.* (2002) précisent que le seul critère texture ne permet pas d'estimer les propriétés de rétention en eau des sols et qu'il faut prendre en compte la nature et le mode d'assemblage des constituants élémentaires, traduit par la densité apparente. On peut citer en outre l'importance de la minéralogie des constituants du sol, les caractéristiques de la roche-mère et la pédogénèse (Bastet, 1999). La minéralogie peut être prise en compte par le biais de la capacité d'échange cationique dans les fonctions de pédotransfert comme le suggère Rawls *et al.* (2001).

Bastet (1999) propose en conséquence une méthodologie d'établissement de fonctions de pédotransfert locales basée sur l'hypothèse que le fait de regrouper les horizons selon des classes peut permettre de rendre communs ces caractères pédo-génétiques et donc fiabiliser la fonction établie pour ces horizons. Les critères de classement utilisés sont la nature du matériau et la texture. Ce qui permet d'améliorer le plus la qualité des fonctions utilisées est la classification par matériau. D'autres auteurs utilisent des approches similaires, Gonçalves *et al.* (1997) utilisent la classe texturale, Pachepsky & Rawls (1999) la classe texturale également plus l'unité taxonomique (matériau) et le régime d'humidité du sol et Wösten *et al.* (2001) rajoutent à cela le comportement du sol. Pachepsky & Rawls (1999) constatent cependant que si la classification permet d'améliorer la qualité des fonctions de pédotransfert sur le jeu de données ayant servi à les établir, il n'en va pas de même sur les données de validation.

Les mesures effectuées, notamment la densité apparente, doivent être réalisées sur des sols non perturbés, comme le soulignent Bruand *et al.* (1995) et Romano & Santini (1997).

De nombreux auteurs concluent sur la relative faiblesse de la qualité de l'estimation des propriétés hydriques des sols via des fonctions ou des règles de pédotransfert (Tietje & Tapkenhinrichs, 1993, Romano & Santini, 1997, Gonçalves *et al.*, 1997, Pachepsky & Rawls, 1999). Mayr & Jarvis (1999) trouvent leurs relations robustes, sauf dans les horizons riches en matière organique, montrant la nécessité de prendre ce facteur en compte. Rawls *et al.* (2001) obtiennent des erreurs de l'ordre de 0.04 $m^3 m^{-3}$ pour la teneur en eau massique des horizons de surface et de 0.065 $m^3 m^{-3}$ pour les horizons profonds à un potentiel de –33 kPa. Schaap *et al.* (2001) obtiennent des valeurs de 0.078 à 0.044 $m^3 m^{-3}$ selon le nombre de facteurs de prédiction pris en compte. Gijsman *et al.* (2002) tirent de la comparaison de plusieurs règles qu'il est difficile de recommander des méthodes par type de sol. L'utilisation de différentes règles pour créer des entrées de modèles conduisent à une variabilité entre les simulations selon eux « inquiétantes ». Une sorte d'analyse de sensibilité des sorties du modèle à la granulométrie montre que la variabilité à l'intérieur d'une même classe texturale est du même ordre de grandeur que la variabilité entre classes. Lin *et al.* (1999) aboutissent à la conclusion que ces valeurs peuvent servir de point de départ à l'estimation des paramètres de sol, constituant ainsi ce qu'on peut appeler une « information *a priori* ».

III.2.2.2.3 Fonctions de pédotransfert, spatialisation et modélisation

Quelle que soit l'échelle d'étude considérée, le renseignement spatial des entrées du modèle passe par l'utilisation de cartes dont les informations sont extrapolées en données utilisables par le modèle par le biais de fonctions de pédotransfert. Gabrielle *et al.* (2002) ont développé une procédure de renseignement des paramètres des entrées sol d'un modèle de culture à partir de propriétés des sols couramment disponibles pour une base de données européenne. Lagacherie *et al.* (2000) utilisent également une carte des sols à l'échelle régionale. Face à la forte variabilité rencontrée à l'intérieur des zones délimitées, ils distinguent plusieurs UTS par UC. Comme les UTS sont caractérisées par des classes texturales, ils utilisent une approche stochastique, ce qui permet d'obtenir des distributions de rendement et d'estimer les erreurs liées à la mauvaise connaissance des paramètres sol en chaque point simulé. Finke *et al.* (1996) et Gijsman *et al.* (2002) ont une approche similaire qui consiste à étudier l'incertitude générée par les fonctions de pédotransfert sur les variables

de sorties de modèles (respectivement, un modèle de fonctionnement hydraulique du sol et un modèle de culture).

Pour effectuer des simulations intra-parcellaire avec STICS, Nicoullaud *et al.* (soumis) ont établi un inventaire cartographique à grande échelle. Une parcelle de 115 ha a fait l'objet d'une cartographie au 1 :50 000 composée de 80 UCS. A chaque UCS est associée une description de tous les horizons constitutifs du sol. Pour chaque horizon, les paramètres nécessaires au modèle sont évalués à l'aide de règles et fonctions de pédotransfert adaptées au milieu d'étude.

Sinowski *et al.* (1997), dans une approche de spatialisation de fonctions de pédotransfert, comparent deux méthodes de spatialisation différentes :

- (i) interpolation des propriétés des sols servant d'entrées à la fonction de pédotransfert et application de celle-ci à ces données interpolées ;

- (ii) application de la fonction de pédotransfert aux propriétés de base des sols puis interpolation des résultats.

Cela revient à se poser la question de savoir s'il est préférable d'interpoler les entrées ou les sorties des fonctions de pédotransfert. La conclusion est que la première approche est la meilleure. Ceci revient donc à utiliser des cartes, qui sont des données spatialisées des entrées des fonctions de pédotransfert et à appliquer ces fonctions aux cartes.

En conclusion, nous pouvons dégager de cette analyse bibliographique les points suivants :

- pour utiliser des données d'entrées sol dans un modèle de culture, il est nécessaire d'établir des cartes des sols à l'échelle adaptée et surtout établir des règles de pédotransfert locales ;

- pour établir ces règles de pédotransfert, il est préférable de grouper les horizons par classes de matériau, texture, origine pédologique, profondeur ;

- il est important de garder à l'esprit l'importante variabilité naturelle des propriétés hydriques des sols au sein des classes de sol que l'on peut constituer.

III.2.3 Présentation des données utilisées

Le modèle STICS a été présenté dans le Chapitre II. Pour réaliser des simulations à l'échelle intra-parcellaire, il est nécessaire, comme on l'a vu, de déterminer des couvertures spatiales de ses différentes variables d'entrée caractéristiques du sol à des échelles de précision compatibles avec l'agriculture de précision. Les mesures effectuées dans ce dessein vont être décrites dans cette partie d'après Nicoullaud *et al.* (2003) et Beaudoin *et al.* (2003).

III.2.3.1 Les variables d'entrée caractéristiques du sol

Le Tableau III.2-2 reprend les différentes entrées caractéristiques du sol de STICS.

Variables	Signification	Acquisition
ARGI	Teneur en argile après décarbonatation de l'horizon de surface (%)	Mesures + krigeage
NORG	Teneur en azote organique de l'horizon de surface (%)	Mesures + krigeage
CALC	Teneur en calcaire de l'horizon de surface (%)	Mesures + krigeage
ALBEDO	*Albedo* de l'horizon de surface	Dérivé de *CALC* et *ARGI*
q0	Seuil de cumul d'évaporation journalière potentielle (mm)	Dérivé de *ARGI*
OBSTARAC	Profondeur du sol à laquelle un obstacle à l'enracinement apparaît (cm)	Carte pédologique + RPT
EPC(h)	Epaisseur de l'horizon *h* décrit par les variables ci-dessous (cm)	Carte pédologique
HCCF(h)	Teneur en eau à la capacité au champ de l'horizon *h* (g eau g^{-1}sol)	Carte pédologique + RPT
HMINF(h)	Teneur en eau au point de flétrissement permanent de l'horizon *h* (g eau g^{-1}sol)	Carte pédologique + RPT
DA(h)	Densité apparente de l'horizon *h* (g cm^{-3})	Carte pédologique + RPT
CAILLOUX(h)	Teneur de l'horizon *h* en cailloux (%)	Carte pédologique + RPT
NO3initf(h)	Azote présent au semis dans l'horizon *h* (kg ha^{-1})	Mesures + krigeage
Hinitf(h)	Teneur en eau au semis dans l'horizon *h* (%)	Mesures + krigeage

Tableau III.2-2 : Présentation des variables d'entrée du modèle STICS décrivant les propriétés du sol et de leurs méthodes d'acquisition respectives. *h* désigne le numéro de l'horizon et prend les valeurs de 1 à 5. Les horizons ont des épaisseurs variables. RPT désigne l'utilisation d'une règle de pédotransfert.

La troisième colonne précise le mode de spatialisation de la variable :

- (i) mesures sur une grille régulière et spatialisation par krigeage ;
- (ii) calcul à partir d'une autre variable mesurée ;
- (iii) obtention à partir d'une carte pédologique descriptive et de règles de pédotransfert permettant de faire le lien entre la description pédologique et la variable.

Les modes (i) et (ii) d'une part et (iii) d'autre part impliquent des méthodes de prospection différentes.

III.2.3.2 Les parcelles expérimentales de Chambry : contexte pédologique

Le site expérimental, on le rappelle, est situé à Chambry, petite commune sise non loin de Laon, dans l'Aisne. Le substrat géologique est la craie du *Sénonien-Campagnien* (ère secondaire). D'autres matériaux viennent s'ajouter à celui-ci :

- les sables du *Thanétien* (ère tertiaire), présents au pied de la butte de Laon ;
- des limons des plateaux, issus de l'érosion éolienne et déposés à l'ère quaternaire.

La parcelle 1, nommée *« Champ du Roi »* au cadastre, est située en position basse et présente une topographie peu accentuée (3 m de dénivelé). Le substrat est composé de craie cryoturbée[1] remaniée par un diluvium crayeux et présentant fréquemment une grève calcaire. Ces formations ont été reprises par des dépôts éoliens de limons et de sables de Sissone. Il résulte de cette histoire géologique complexe une forte hétérogénéité spatiale des dépôts, parfois d'ordre métrique.

La parcelle 2, répondant au nom de *« Grande Couture »*, se situe en position haute sur la plaine et présente une topographie plus nette (7 m de dénivelé). Le substrat est composé de craie blanche avec un niveau supérieur de craie sableuse magnésienne. La craie a également subi une cryoturbation et peut être recouverte de sables du Thanétien et de limons éoliens. Le sommet de la parcelle et le versant sud-sud-ouest, soumis à l'érosion éolienne, sont occupés par des sols calcaires. En revanche, le versant opposé est formé de sols profonds limoneux et moins calcaires.

---

[1] Cryoturbation : modification du sol sous l'effet du gel, par gélivation, soulèvement, etc.

III.2.3.3 Mesures de caractéristiques physico-chimiques de l'horizon de surface

Les variables suivantes : *NORG*, *ARGI* et *CALC* ont été mesurées sur les nœuds d'une grille systématique espacés de 36 m. La parcelle 1 en possède 81 et la parcelle 2, 85 (*cf.* Figure III.2-2 pour la parcelle 1 et Figure III.2-3 pour la parcelle 2). Des mesures réalisées sur des croix, avec un pas de 2 m, ont complété le dispositif de mesure afin de caractériser la variabilité des grandeurs mesurées à courte distance et permettre l'élaboration et la modélisation de semi-variogrammes utilisés pour le krigeage.

ARGI correspond dans STICS à la teneur en argile après décarbonatation. Or, seules les teneurs en argile granulométrique, sans décarbonatation, ont été mesurées exhaustivement sur la grille. Cependant, les teneurs en argile avant et après décarbonatation ont été déterminées sur une cinquantaine d'échantillons, ce qui a permis d'établir une relation empirique locale entre teneur en argile avant décarbonatation, teneur en calcaire et *ARGI*. Grâce à cette relation, *ARGI* a pu être évaluée sur l'ensemble de la grille (Beaudoin *et al.*, 2003).

III.2.3.4 Réalisation de sondages pédologiques sur les grilles

Sur tous les nœuds des grilles précédemment définies, des sondages ont été réalisés manuellement à la tarière jusqu'à 1.2 m. Les différents horizons ont été décrits à dire d'expert (Legros *et al.*, 1996) par l'estimation de différentes caractéristiques : profondeur, couleur, texture par détermination tactile, teneur en calcaire, pierrosité, degré d'hydromorphie, compacité. Les prélèvements effectués au carottier mécanique pour effectuer les mesures de teneur en eau et en azote jusqu'à 1.5 m (*cf. infra*) ont également été utilisés en complément. En outre, compte tenu de l'hétérogénéité de la parcelle, des sondages supplémentaires ont été ajoutés en se guidant sur la topographie établie au tachéomètre et sur des images aériennes de la parcelle. Le système utilisé alliait donc une prospection systématique et une prospection orientée par d'autres observations. La succession et la caractérisation des horizons pédologiques ont ainsi été déterminées sur tous les nœuds de la grille et ont permis d'établir la carte des sols.

III.2.3.5 Caractérisation de propriétés du sol sur des fosses pédologiques

Quatre à cinq fosses pédologiques ont été creusées sur chaque parcelle. Elles ont fait l'objet de diverses mesures et observations :

- description des horizons d'après la méthode décrite par Gaultier *et al.* (1993) ;

- mesures sur des prélèvements des teneurs en eau au point de flétrissement permanent en laboratoire ;

- mesures de densité apparente ;

- mesures de propriétés physico-chimiques : granulométrie, teneur en calcaire ;

- établissement de cartes d'impacts racinaires selon la méthode décrite par Tardieu et Manichon (1986).

Ces observations permettront d'établir les règles de pédotransfert concernant les densités apparentes et les humidités au point de flétrissement permanent d'une part et l'enracinement d'autre part.

III.2.3.6 Mesures de la teneur en eau et des reliquats d'azote

Des mesures des caractéristiques non permanentes de la parcelle ont été régulièrement réalisées sur les nœuds de la grille. Il s'agissait de la teneur en eau et des reliquats azotés. Les prélèvements étaient réalisés grâce à un carotteur mécanique, par tranches de 30 cm jusqu'à 150 cm dans la mesure du possible. L'azote minéral était extrait par une solution de chlorure de potassium (KCl) molaire (ratio sol:solution = 1:5). Les mesures de teneurs en nitrates étaient réalisées à l'aide d'un analyseur TRAACS 2000 (Bran & Luebbe, Allemagne) en utilisant la méthode proposée par Kamphake *et al.* (1967).

Le Tableau III.2-3 présente le protocole et les dates de mesures pour les deux parcelles pour les années 1999 à 2003. Nous ne traiterons que les mesures concernant la culture du blé.

Les croix servent à étudier la variabilité à faible distance des mesures et seront utilisées pour réaliser les krigeages de ces mesures (établissement et modélisation des semi-variogrammes). Le dispositif sur la parcelle 1 est assez peu équilibré. Au semis du blé de 2001 par exemple, seuls quatre points de mesure ont été effectués. Mais après une culture de betterave comme c'était le cas en 2001, les reliquats azotés sont faibles et par conséquent assez peu variables géographiquement (Machet, comm. pers.). La moyenne des mesures obtenues sur les quatre points sera utilisée sur toute la parcelle.

Parcelle 1			Parcelle 2		
Date et phase	Grille	Croix	Date et phase	Grille	Croix
05/10/99 : semis du blé	22		04/10/00 : semis du blé	85	
21/02/00 : tallage du blé	81	4	14/02/01 : tallage du blé	85	4
16/08/00 : récolte du blé	22		01/08/01 : récolte du blé	85	
04/10/01 : semis du blé	4		30/09/02 : semis du blé	85	
30/01/02 : tallage du blé	81	4	12/02/03 : tallage du blé	85	4
22/07/02 : récolte du blé	81				

Tableau III.2-3 : Mesures de teneur en eau et de reliquats en azote sur les deux parcelles expérimentales de Chambry. La colonne **Grille** indique le nombre de nœuds sur lesquels ont été effectuées des mesures. Les nœuds de la grille étaient régulièrement espacés de 36 m et positionnés en quinconce. La colonne **Croix** donne le nombre de points où les mesures ont été réalisées dans les quatre directions cardinales sur 24 m avec un pas de 2 m. Les mesures effectuées à la récolte du blé en 2003 ne sont pas encore traitées.

III.2.4 Caractérisation spatiale des entrées sol de STICS

La méthode employée pour caractériser l'ensemble des variables d'entrée sol de STICS reprend les différents aspects vus dans l'analyse bibliographique. La méthode de spatialisation employée dépend de la nature de la variable (Tableau III.2-2).

III.2.4.1 Etablissement des cartes pédologiques à grande échelle

Ce travail a été réalisé par l'INRA d'Orléans (B. Nicoullaud, J. Roque et A. Couturier) en collaboration avec l'INRA de Laon (N. Beaudoin). La description concise qui va en être faite se base sur ce qui est rapporté par Nicoullaud *et al.* (2003).

Les contours des Unités Cartographiques de Sol ont été établis à partir des observations ponctuelles, de la topographie et des observations aériennes des parcelles en sol nu. A cette base de données géographique est associée une base de données sémantique qui décrit la proportion de surface de chaque UCS occupée par une ou plusieurs Unités Typologiques de Sol. Chaque horizon de celles-ci est décrit jusqu'à 1.5 m de profondeur au plus par des caractéristiques (i) qualitatives, (ii) quantitatives et (iii) synthétiques, respectivement :

- classe texturale (par exemple, « LAS » pour Limon Argilo-Sableux) en intégrant pour les horizons profonds – en deçà de 30 cm – des notions de couleur et de compacité susceptibles de correspondre à des propriétés différentes comme la densité apparente. Cette classe texturale intègre dans ce cas en même temps pour les horizons profonds

une notion de structure liée au matériau parental ou à l'évolution pédologique. On utilise dans ce cas là le terme de classe texturo-structurale ;

- profondeur, teneur en calcaire, en cailloux et en graviers ;
- substrat, type de profil, drainage interne.

Le même type de base de données a été établi pour les nœuds des grilles de mesure.

Les cartes obtenues sont présentées respectivement dans la Figure III.2-2 pour la parcelle 1 et dans la Figure III.2-3 pour la parcelle 2.

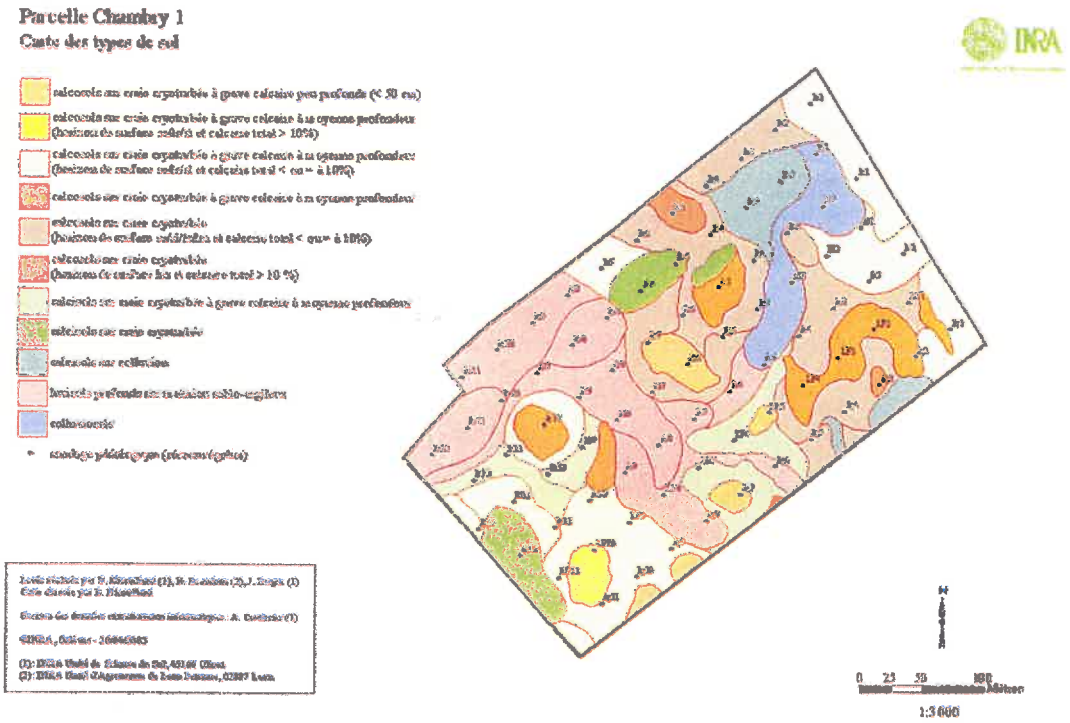


Figure III.2-2 : Carte pédologique au 1:3000 de la parcelle 1. Extrait de Nicoullaud *et al.*, 2003. Les points figurent les nœuds de la grille régulière de maille 36 m.

Pour la parcelle 1, 50 unités cartographiques ont été différenciées et ont pu être regroupées en 11 grands types de sols. Une grande variabilité à l'échelle métrique est observée : celle-ci pourrait être liée à l'apport de différents matériaux par dépôts éoliens et à la conservation des poches de cryoturbation expliquée par une érosion moindre que dans le cas de la parcelle 2. Par ailleurs, la zone a été le théâtre de combats au cours des deux guerres mondiales.

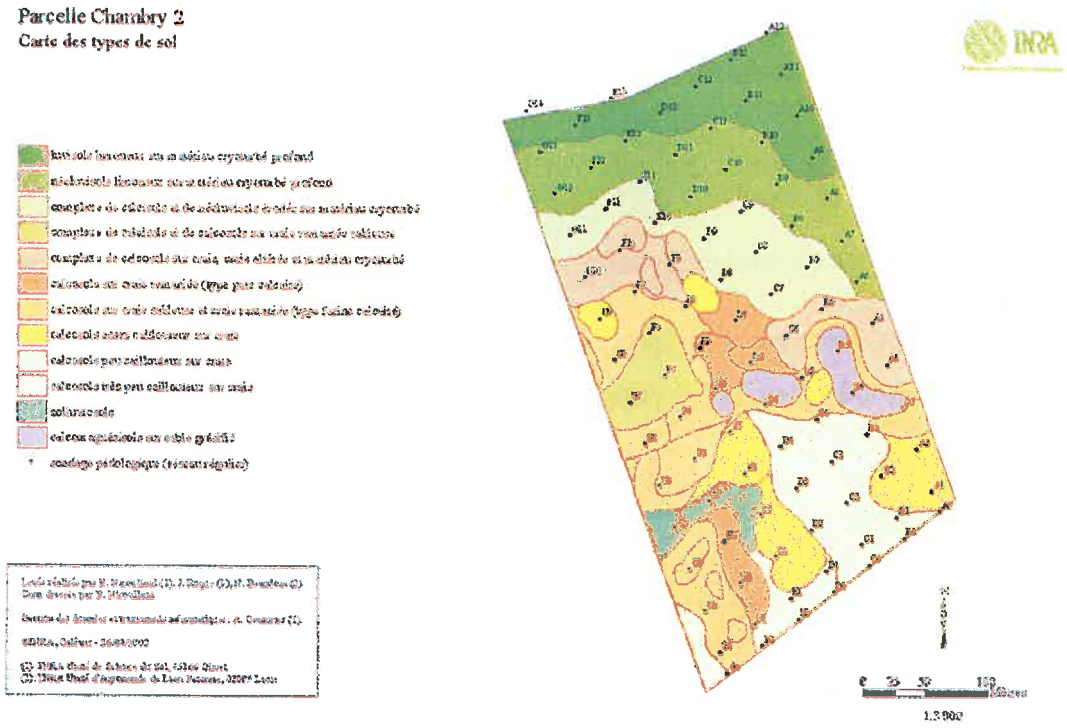


Figure III.2-3 : Carte pédologique au 1:3000 de la parcelle 2. Extrait de Nicoullaud *et al.*, 2003. Les points figurent les nœuds de la grille régulière de maille 36 m.

Sur la parcelle 2, 42 unités cartographiques ont été différenciées, regroupées en 12 grands types de sol. La variabilité à l'échelle métrique y est très faible, au moins dans les premiers horizons. La topographie de la parcelle est plus marquée ; sa géomorphologie, majoritairement convexe, et sa localisation, en position sommitale en haut de versants, ont favorisé l'érosion. Les matériaux actuels sont hérités pour leur plus grande part des substrats géologiques, plus homogènes géographiquement. Des restes de poches de cryoturbation génèrent cependant une variabilité d'ordre *infra*-métrique seulement en profondeur.

La topographie plus marquée et caractérisée par une plus grande convexité que dans la parcelle 1 est à l'origine d'une moindre variabilité géographique. En effet, ce type de conformation est plus soumis à l'érosion, ce qui fait que l'héritage des roches mères, elles-mêmes moins variables, se fait dans ce cas davantage sentir.

Une base de données similaire à celle qui a été constituée sur les Unités Cartographiques de Sol a été établie sur les nœuds des grilles : on possède donc deux référentiels géographiques différents sur lesquels les simulations pourront être réalisées.

III.2.4.2 Constitution des règles de pédotransfert

Cet aspect a été développé en collaboration avec B. Nicoullaud (INRA Orléans) et N. Beaudoin (INRA Laon).

Les règles de pédotransfert effectuées ont été établies sur des supports différents selon les variables traitées. Elles concernent les teneurs en eau à la capacité au champ, les teneurs en eau au point de flétrissement permanent, les densités apparentes et les propriétés vis-à-vis de l'enracinement. Les noms de ces variables seront abrégés dans la suite d'après le nom de la variable de STICS correspondante, soit, respectivement : *HCCF(h)*, *HMINF(h)*, *DA(h)* et *OBSTARAC*.

III.2.4.2.1 Règle de pédotransfert concernant *HCCF(h)*

Les règles de pédotransfert (RPT) concernant cette variable ont comme entrées des types d'horizon, c'est-à-dire la combinaison entre la classe texturale pour les horizons de surface et la classe texturo-structurale pour les horizons profonds, sa profondeur et sa teneur en calcaire. Ces entrées correspondent au contenu de la base de données caractérisant chaque type d'UTS.

Elles ont été établies au niveau des nœuds des grilles de mesures systématiques. Les teneurs en eau mesurées en février (*cf.* Tableau III.2-3) ont été utilisées, en posant l'hypothèse qu'elles représentaient la teneur en eau à la capacité au champ. Ces mesures étaient en effet effectuées au cours de l'hiver, après des pluies et un temps suffisamment long pour permettre le ressuyage. Les mesures du 21 février 2000 et du 30 janvier 2002 ont été utilisées pour la parcelle 1 et celles du 20 mars 2000 et du 14 février 2001 pour la parcelle 2. Les deux années de mesures ont permis de vérifier cette hypothèse. Sur la parcelle 1, la corrélation linéaire entre les mesures de teneur en eau en sortie d'hiver des deux années de mesures est de R^2=0.77 avec une droite de régression proche de la première bissectrice. Pour la parcelle 2, la corrélation est de R^2=0.82, la droite de régression étant proche également de la première bissectrice. L'hypothèse selon laquelle les mesures en sortie d'hiver peuvent être assimilées à la teneur à la capacité au champ n'est donc pas invalidée, bien que la relative dispersion doive être considérée comme une source d'erreur possible.

Parmi toutes les carottes de 30 cm prélevées et dont la teneur en eau en février a été mesurée, celles qui étaient à cheval sur deux horizons, dites carottes hétérogènes, ont été supprimées de la base de données. Elles serviront à valider les règles de pédotransfert

concernant *HCCF(h)* et *DA(h)*. Seules les carottes homogènes ont servi à établir la règle de pédotransfert (Figure III.2-4).

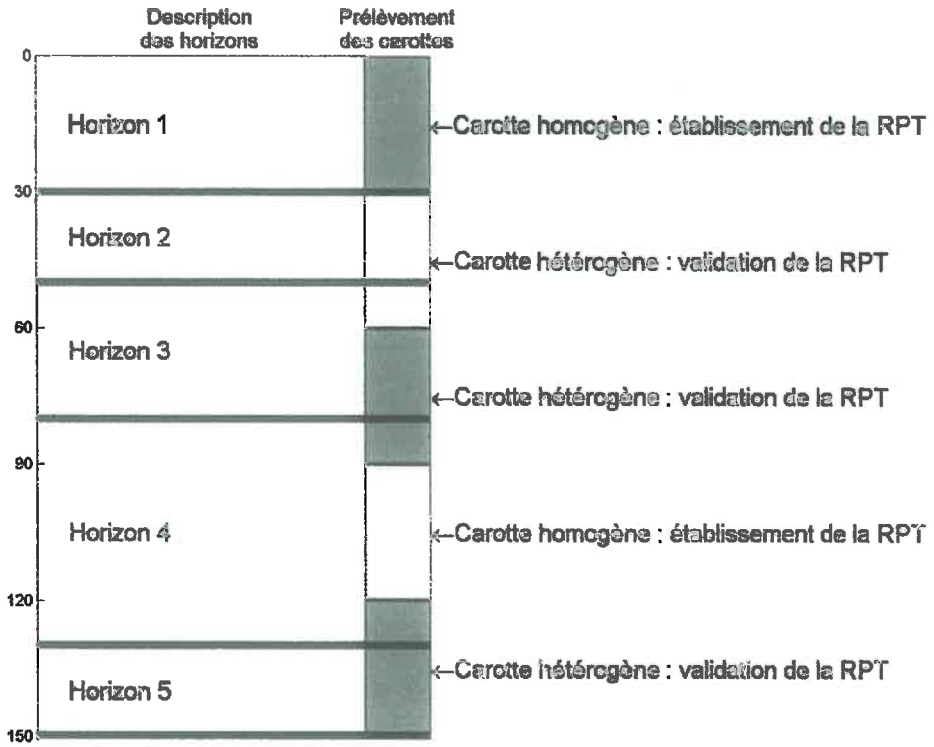


Figure III.2-4 : Représentation schématique de la différenciation entre les carottes de 30 cm utilisées pour établir la règle de pédotransfert sur *HCCF(h)*, dites carottes homogènes, et celles qui ont été utilisées pour la valider, dites carottes hétérogènes.

Les différentes carottes ont été regroupées par classes selon les critères suivants :

- type d'horizon (de surface ou profond) ;
- classe texturale ou classe texturo-structurale ;
- teneur en calcaire.

Ces critères constitueront les entrées de la règle de pédotransfert. La moyenne des carottes appartenant à une classe donnée est attribuée à cette classe. Les résultats de cette classification sont donnés sous forme graphique dans la Figure III.2-5. Les valeurs numériques sont indiquées selon les classes de pédotransfert, dans le Tableau III.2-4 et le Tableau III.2-5 respectivement pour la parcelle 1 et 2.

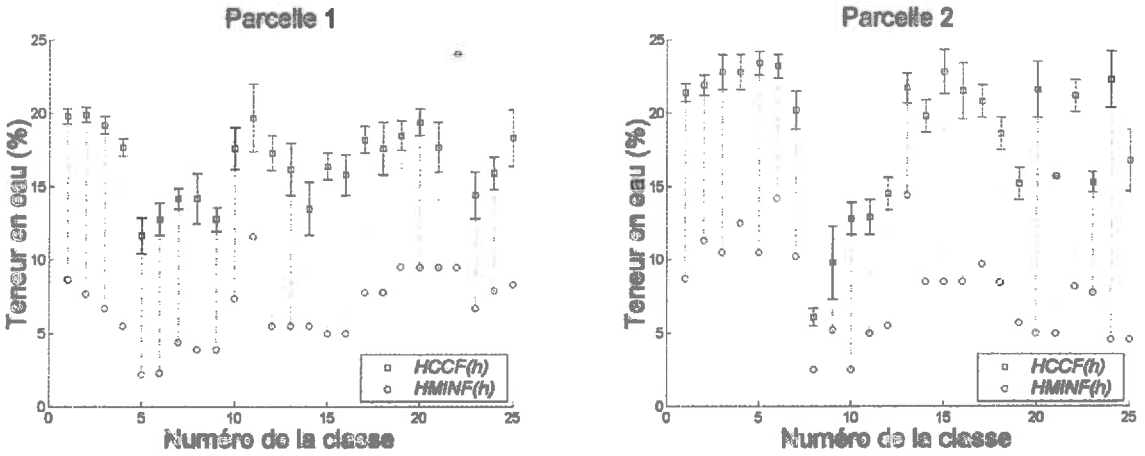


Figure III.2-5 : Représentation, pour les parcelles 1 et 2, des valeurs de *HCCF(h)* (avec les écarts types) et de *HMINF(h)* (écarts types non calculables) en fonction des classes. Les traits pointillés servent à associer les valeurs correspondantes de *HCCF(h)* et de *HMINF(h)*.

On observe une variabilité intra-classe relativement forte comme en témoignent les écarts types figurés pour *HCCF(h)* alors que la variabilité inter-classe n'est pas toujours très importante ; ceci est à rapprocher des inquiétudes de Gijsman *et al.* (2002) qui observaient semblable phénomène.

III.2.4.2.2 Règles de pédotransfert concernant *HMINF(h)* et *DA(h)*

Pour toutes les classes d'horizon définies ci-dessus et retranscrites dans le Tableau III.2-4 et le Tableau III.2-5, des mesures de *HMINF(h)* et de *DA(h)* ont été réalisées sur les horizons des fosses pédologiques, assez représentatives de la variabilité des parcelles. Certains horizons rencontrés sur les parcelles mais pas dans les fosses ont fait l'objet de mesures supplémentaires. Pour ces deux variables, une seule mesure a été effectuée par classe d'horizon. Les résultats sont donnés dans la Figure III.2-5 pour *HMINF(h)* et dans le Tableau III.2-4 et le Tableau III.2-5 respectivement pour les parcelles 1 et 2 pour les deux variables.

III.2.4.2.3 Validation des règles de pédotransfert concernant *HCCF(h)* et *DA(h)*

Afin de donner une validation des règles de pédotransfert concernant les *HCCF(h)* et les *DA(h)*, nous avons utilisé les carottes hétérogènes. Les valeurs de *HCCF(h)* et de *DA(h)* ont été déterminées pour chaque fragment de carotte homogène ; une valeur de teneur en eau à la capacité au champ de la carotte a été calculée en utilisant la moyenne des *HCCF(h)* pondérée par le produit des *DA(h)* et des épaisseurs d'horizons homogènes dans chaque carotte (Figure III.2-6). Cette valeur est ensuite comparée à la valeur mesurée.

Texture et taux de calcaire	Couleur	Profond. (cm)	*HCCF* (%)	*HMINF* (%)	*DA* ($g\ cm^{-3}$)
LSA		0 – 30	19.8±0.5 ; 9	8.6	1.41
LS		0 – 30	19.9±0.5 ; 19	7.7	1.47
SA		0 – 30	19.2±0.6 ; 45	6.7	1.46
S		0 – 30	17.7±0.6 ; 3	5.5	1.48
SL		0 – 30	*Id*	*Id*	1.45
S	Jaune, rouge, brun	30 – 150	11.6±1.2 ; 21	2.2	1.58
Id	Vert, vert-jaune	30 – 150	12.8±1.1 ; 5	2.3	1.58
SL et K<10	Brun clair, gris, jaune	30 – 150	14.2±0.7 ; 10	4.4	1.55
SL et K>10	Brun clair, gris, jaune	30 – 90	14.2±1.7 ; 4	3.9	*Id*
Id	*Id*	90 – 150	12.8±0.8 ; 13	*Id*	*Id*
SA	Brun, rouge	30 – 150	17.6±1.4 ; 9	7.4	1.56
LSA	Brun	30 – 150	*Id*	*Id*	1.51
A, AS	Brun, brun rouge	30 – 150	19.7±2.3 ; 8	11.6	1.59
SA et K>10	Beige, brun, ocre-gris, vert	30 – 90	17.3±1.2 ; 14	5.5	1.56
Id	*Id*	90 – 120	16.2±1.8 ; 4	*Id*	*Id*
Id	*Id*	120 – 150	13.5±1.8 ; 12	*Id*	*Id*
SA et K<10	Beige, jaune, brun	30 – 90	16.4±0.9 ; 39	*Id*	*Id*
SA, LS et K<10	*Id*	90 – 150	15.8±1.4 ; 5	*Id*	*Id*
LSA et K>10	Beige, brun-jaune, ocre	30 – 150	18.2±0.9 ; 4	7.8	1.51
LSA et K<10	Beige, brun-jaune clair	30 – 150	17.6±1.8 ; 13	*Id*	*Id*
LAS et K>10	Beige	30 – 150	18.5±1.0 ; 14	9.5	1.50
Id	Ocre-gris, brun-clair, jaune	30 – 150	19.4±0.9 ; 9	*Id*	*Id*
Id	Gris, gris brun, gris clair	30 – 150	17.7±1.7 ; 6	*Id*	*Id*
Id	Brun foncé, gris	30 – 150	24.0 ; 2	*Id*	*Id*
LS* et K>10	Gris	30 – 150	14.4±1.6 ; 16	6.7	1.66
LAS* et K>10	*Id*	30 – 150	*Id*	*Id*	1.72
LAS* et K>10	Gris	30 – 150	15.9±1.1 ; 12	7.9	1.65
LAS et K<10	Brun jaune, gris	30 – 150	18.3±1.9 ; 8	8.3	1.50

Tableau III.2-4 : Classes de pédotransfert indiquant pour la parcelle 1 les valeurs de *HCCF(h)*, de *HMINF(h)* et de *DA(h)* en fonction de la texture du matériau de l'horizon, de sa teneur en calcaire (K), de sa couleur et de sa profondeur. Les symboles des textures font référence à ceux qui sont utilisés dans le triangle des textures de l'Aisne. Le symbole ± précède l'écart type calculé sur l'effectif donné après le point virgule. Le symbole * indique une compacité sensible à la tarrière. *Id* indique que la statistique est commune avec la ligne du dessus.

Texture et taux de calcaire	Couleur	Profond. (cm)	*HCCF (%)*	*HMINF (%)*	*DA ($g\ cm^{-3}$)*
LMS, LAS^-		0 – 30	21.4±0.6 ; 7	8.7	1.40
LAS et K<10		0 – 30	21.9±0.7 ; 11	11.3	1.45
LAS et K>10		0 – 30	22.8±1.2 ; 30	10.5	1.35
LAS^+ et K<10		0 – 30	*Id*	12.5	1.40
LAS^+ et K>10		0 – 30	23.4±0.8 ; 15	10.5	1.35
A		0 – 30	23.2±0.8 ; 13	14.2	1.40
LSA		0 – 30	20.2±1.3 ; 13	10.2	1.45
Grès		30 – 150	6.1±0.6 ; 10	2.5	1.90
S Tf		30 – 150	9.8±2.5 ; 9	5.2	1.60
S	Vert, jaune	30 – 150	12.8±1.1 ; 5	2.5	1.50
SL	Brun-jaune	30 – 150	12.9±1.2 ; 6	5	1.50
SA	Ocre-jaune, jaune	30 – 150	14.5±1.1 ; 7	5.5	1.50
AS et K>10	Brun-jaune	30 – 150	*Id*	*Id*	*Id*
A	Brun-rouge, brun-ocre	30 – 150	21.7±1.0 ; 9	14.0	1.50
AL	Brun-jaune	30 – 150	*Id*	*Id*	*Id*
AS et K<10	Brun-jaune	30 – 150	*Id*	*Id*	*Id*
LAS et K<10	Brun-jaune, rouge	30 – 150	19.8±1.1 ; 11	8.5	1.45
LAS et K>10	Beige, brun	30 – 150	22.9±1.5 ; 21	8.5	1.45
LAS sur craie	Beige, brun	30 – 150	21.5±1.9 ; 15	8.5	1.45
LAS et K>10	Jaune-ocre, ocre-jaune	30 – 150	20.8±1.1 ; 6	9.7	1.45
LSA et K>10	Beige, brun-jaune foncé	30 – 150	18.6±1.1 ; 3	8.4	1.60
LS	Beige, brun-jaune foncé	30 – 150	15.2±1.1 ; 8	5.7	1.60
Craie	Blanche	30 – 150	21.6±1.9 ; 49	5.0	1.45
Craie	Grise	30 – 150	15.7 ; 2	5.0	1.60
Craie alt.	Blanche	30 – 150	21.2±1.1 ; 6	8.2	1.50
Craie alt.	Jaune-ocre	30 – 150	15.3±0.7 ; 8	7.8	1.60
LMS	Beige	30 – 150	22.3±1.9 ; 15	4.6	1.50
LS		30 – 150	16.8±2.1 ; 18	4.6	1.60

Tableau III.2-5 : Classes de pédotransfert indiquant pour la parcelle 2 les valeurs de *HCCF(h)*, de *HMINF(h)* et de *DA(h)* en fonction de la texture du matériau de l'horizon, de sa teneur en calcaire (K), de sa couleur et de sa profondeur. Les symboles des textures font référence à ceux qui sont utilisés dans le triangle des textures de l'Aisne. Tf : très fin ; alt. : altéré. Le symbole ± précède l'écart type calculé sur l'effectif donné après le point virgule. *Id* indique que la statistique est commune avec la ligne du dessus.

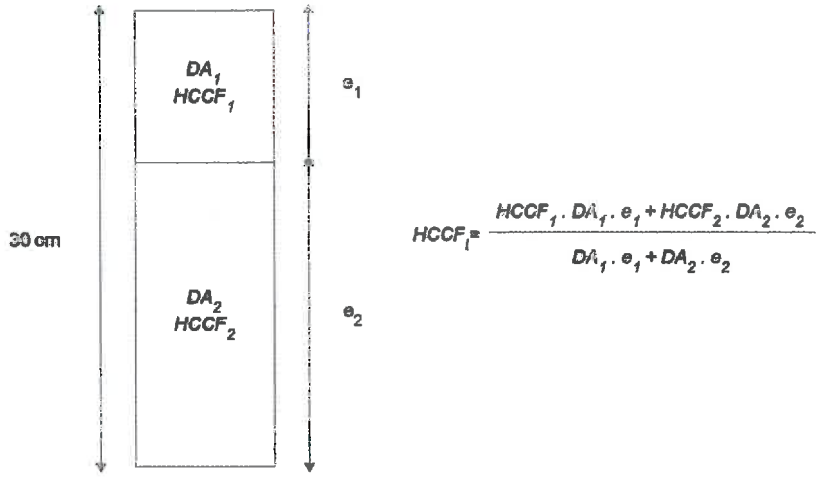


Figure III.2-6 : Calcul de la teneur en eau à la capacité au champ sur les carottes hétérogènes d'après les propriétés déterminées par les règles de pédotransfert relatives à *HCCF(h)* (g d'eau g^{-1} de sol) et *DA(h)* (g cm^{-3}).

Les résultats de la confrontation sont, en terme de RMSE, de 1.61 pour la parcelle 1 et de 1.90 pour la parcelle 2. Ces résultats sont inférieurs aux valeurs annoncées par Rawls *et al.* (2001), Schaap *et al.* (2001) et Bruand *et al.* (2002) (*cf. supra*), ce qui justifie le travail d'élaboration de règles locales.

III.2.4.2.4 Règle de pédotransfert concernant *OBSTARAC*

Contrairement aux règles de pédotransfert précédentes, celle-ci a comme entrée un profil et non un horizon particulier caractérisé par sa profondeur, son matériau et sa teneur en calcaire.

Sur les fosses pédologiques, la succession des types d'horizons dans le profil a été mise en rapport avec l'enracinement maximal. Les règles de décision suivantes, pour décider de la profondeur maximale d'enracinement pour un profil, ont été établies. On distingue deux types d'horizons susceptibles de stopper l'enracinement : (i) les horizons d'arrêt 'probable' sont seulement suspectés de stopper l'enracinement ; (ii) les horizons d'arrêt 'certain' sont supposés représenter un obstacle irrécusable à l'enracinement.

- Les horizons d'arrêt probable sont : les horizons compacts (marqués d'une * dans les tableaux donnant les classes de pédotransfert) et LAS, LS, LA, SA très calcaires (K>40).

- Les horizons d'arrêt certain sont : S, SL, LS à graviers, calcaire, craie et, pour une profondeur supérieure à 90 cm, tout matériau à teneur en graviers supérieure à 20%.

La valeur d'*OBSTARAC*, pour un profil donné, est fixée entre le milieu du premier horizon d'arrêt probable rencontré et le début d'un horizon d'arrêt certain. S'il n'y a pas d'horizon d'arrêt probable dans le profil avant le substrat, *OBSTARAC* correspond à l'horizon d'arrêt certain, qui peut être le substrat lui-même.

III.2.4.3 Constitution de cartes krigées et calcul de *ALBEDO* et *q0*.

Les mesures de teneur en eau et en nitrates aux différentes dates par horizon de 30 cm (*cf.* Tableau III.2-3) ont été krigées (Bruchou et Pringle, comm. pers.).

Les cartes obtenues permettent de renseigner les valeurs initiales de teneur en eau et en azote de STICS au semis et de valider les simulations faites à la récolte.

Les valeurs de *ALBEDO* et de *q0* ont été établies respectivement en fonction de *CALC+ARGI* et de *ARGI* selon des tableaux de correspondance présentés dans la notice de STICS 1.0 (Beaudoin, comm. pers.) et reproduits ici :

Type de sol	Albédo	Argile (%)	Limon (%)	Sable (%)	CaCO3 (%)
Craie	0.31	25	52	24	52
Sablo-Limoneux	0.25	8	20	72	9
Limon Argileux	0.18	24	47	29	7
Limon Moyen	0.22	12	79	7	3
Limon Moyen battant	0.28	12	79	7	3
Limon Moyen 2	0.23	16	79	4	3
Limon Argileux	0.22	23	72	5	3

Tableau III.2-6 : Table de correspondance entre le type de sol et sa teneur en calcaire et l'*albedo* (extrait de la notice de STICS 1).

Type de sol	q0
Argile < 5 %	6
5 % < Argile < 25 %	9
25 % < Argile < 45 %	12
Argile > 45 %	6

Tableau III.2-7 : Table de correspondance entre le type de sol et la valeur de *q0* (mm), extrait de la notice de STICS 1.

III.2.5 Conclusion

L'ensemble des variables d'entrée concernant le sol de STICS présentées dans le Tableau III.2-2 ont donc fait l'objet de mesures qui peuvent être spatialisées à l'échelle de la parcelle par des méthodes différentes.

Une base de données sémantique décrivant les successions d'horizons a été établie à deux échelles : (i) les nœuds de mesure de la grille ; (ii) les Unités Cartographiques de Sol. Le contenu de cette base, la description de la succession des horizons, sert d'entrée à un ensemble de règles de pédotransfert qui permettent de déterminer, pour l'entité géographique donnée, les variables d'entrée sol de STICS. Par ailleurs, les données initiales, contenu en eau et en azote au semis, ont été mesurées à chaque début de campagne et spatialisées par krigeage.

Les sources d'erreur de cette méthode se rencontrent à plusieurs niveaux :

- lissage de la variabilité intra-parcellaire par regroupement de zones ;
- mauvaise caractérisation et description des horizons par diagnostic tactile ;
- erreurs générées par le passage des caractérisations de profils aux variables d'entrée du modèle par l'utilisation des règles de pédotransfert ;
- règles de pédotransfert elles-mêmes.

Ces erreurs se rajouteront aux erreurs de structure du modèle qui font que, pour un système parfaitement décrit en entrée, on n'a pas forcément une bonne simulation des phénomènes.

Cet ensemble d'outils sera utilisé dans le Chapitre IV pour réaliser des simulations à l'échelle intra-parcellaire. Cette approche sera comparée à celle qui est basée sur l'utilisation de l'assimilation de mesures issues de la télédétection et dont le développement fait l'objet de la partie suivante.

III.3 Spatialisation du modèle par assimilation de données

Cette partie est consacrée à la mise en place d'un système de spatialisation des simulations de STICS dans le cadre de l'agriculture de précision basé sur l'assimilation. Trois méthodes seront comparées en vue d'en retenir une qui sera utilisée sur les données de Chambry dans le Chapitre IV. Le jeu de données impliqué dans toute cette partie méthodologique ayant été créé par simulation, nous ne présenterons pas ici la façon dont sont obtenues les données issues de la télédétection servant à l'assimilation, une partie du chapitre suivant y étant consacré. Mais afin que cette partie puisse être lue de façon indépendante, l'introduction qui suit permet de resituer cette étude dans le cadre de notre travail.

III.3.1 Objectif et principe de l'assimilation de données issues de la télédétection dans le cadre de l'agriculture de précision

L'objectif de l'assimilation de données est de caractériser l'état du couvert temporellement et à l'échelle intra-parcellaire. Il s'agit donc bien d'une alternative à la simulation basée sur une approche par cartographie des sols développée dans la partie précédente. Nous avons choisi dans le Chapitre I d'utiliser des méthodes d'assimilation qui permettent d'estimer les variables d'entrée du modèle, de préférence aux méthodes qui corrigent les trajectoires des variables simulées. Or, les données issues de la télédétection se présentent sous forme de cartes, donnant des valeurs de variables biophysiques sur des pixels dont la taille varie selon les capteurs. Il s'agit donc, dans cette approche, de déterminer les caractéristiques du sol servant de variables d'entrée au modèle à l'échelle de ces pixels.

III.3.1.1 Données disponibles et variables du modèle à estimer

Dans l'étude qui suit, nous supposerons acquis que les données accessibles par télédétection et utilisées pour l'assimilation dans le modèle sont l'indice de surface foliaire, *LAI*, et la quantité d'azote absorbée par la culture, *QN*. Nous verrons dans le Chapitre IV comment ces données sont évaluées à l'échelle intra-parcellaire sur la base d'un cas réel.

Ces informations sont donc connues sous la forme de pixels. Pour un capteur de résolution 50×50 m^2 (SPOT), une parcelle de 10 ha compte 40 pixels ; pour un capteur de résolution 2×2 m^2 (CASI), la même parcelle en comptera 25 000. Cela fait autant d'entités sur lesquelles les variables du modèle devront être estimées. Dans toute notre étude, cette estimation se fera de façon indépendante, ce qui sera commenté dans la discussion générale.

Trois méthodes d'assimilation seront comparées dans cette partie méthodologique. Nous considèrerons pour ce faire une parcelle hypothétique formée de 20 pixels seulement dont nous chercherons à estimer les caractéristiques du sol. Néanmoins, des études préalables ayant montré leur importance, trois paramètres caractéristiques de la plante seront ajoutés, en posant l'hypothèse que leur valeur peut être la traduction d'interaction sol-plante non prises en compte par le modèle. Sous cette hypothèse, ils seront eux aussi apparentés à des variables, au sens géographique du terme : il s'agit de $vlai_{max}$, de *durvieF* et de *lvfront*. L'ensemble des variables candidates est résumé dans le Tableau III.3-1. La variable *PROFHUM* n'a pas été considérée, car, avec *NORG*, elle contribue à définir une quantité d'azote organique disponible. Leurs effets sont donc strictement similaires. Elle a été fixée à la valeur de 30 cm, à savoir la profondeur du premier horizon. Les cailloux n'ont pas non plus été pris en considération, leur effet sur la réserve utile pouvant être englobés dans les variables décrivant la réserve utile propre au sol.

Variables	Signification
ARGI	Teneur en argile de l'horizon de surface (%)
NORG	Teneur en azote organique de l'horizon de surface (%)
CALC	Teneur en calcaire de l'horizon de surface (%)
ALBEDO	Albédo de l'horizon de surface
q0	Seuil de cumul d'évaporation journalière potentielle (mm)
OBSTARAC	Profondeur du sol à laquelle un obstacle à l'enracinement apparaît (cm)
HCCF(h)	Teneur en eau à la capacité au champ de l'horizon *h* (g eau g^{-1}sol)
HMINF(h)	Teneur en eau au point de flétrissement permanent de l'horizon *h* (g eau g^{-1}sol)
DA(h)	Densité apparente de l'horizon *h* (g cm^{-3})
NO3initf(h)	Azote présent au semis dans l'horizon *h* (kg ha^{-1})
lvfront	Paramètre plante de croissance racinaire (cm cm^{-3})
durvieF	Paramètre plante de durée de vie des feuilles
$vlai_{max}$	Paramètre plante d'accroissement du *LAI*

Tableau III.3-1 : Présentation des variables d'entrée du modèle STICS décrivant les propriétés du sol et dont une partie sera estimée par assimilation. *h* désigne le numéro de l'horizon et prend les valeurs de 1 à 5.

III.3.1.2 Principe général de l'assimilation

Toutes les variables présentées ci-dessus ne seront cependant pas prises en compte dans l'assimilation. En effet, toutes n'ont pas la même importance dans les simulations et certaines d'entre elles ont des effets compensés. Une première étape consistera donc à réaliser une analyse de sensibilité destinée à sélectionner un certain nombre de variables parmi les plus importantes.

La deuxième étape consiste en l'assimilation de données proprement dite. Cette utilisation de données issues de la télédétection pour estimer les entrées de modèles de culture devient une technique classique. Dans une revue bibliographique, Moulin *et al.* (1998) montrent que les stratégies d'assimilation appliquées en général passent par la minimisation de la différence entre observations et simulations en ajustant un faible nombre de variables d'entrée. En effet, lorsqu'on n'introduit pas d'information *a priori* (IAP) sur la valeur des variables que l'on cherche à estimer, le nombre de variables d'entrée estimées doit être limité en fonction du nombre des observations. Guérif & Duke (2000) et Guérif & Launay (2003) estiment un nombre limité de variables d'entrée (5 et 4 respectivement) du modèle de fonctionnement SUCROS couplé au modèle de transfert radiatif SAIL (pour SAIL, *cf. infra*) en utilisant un algorithme de minimisation de la différence entre les valeurs observées et simulées des réflectances dérivé du simplexe. Dans ces travaux, des problèmes de compensation d'effet entre différentes variables sont rencontrés. Prévot *et al.* (2003) utilisent quant à eux le modèle STICS couplé au modèle de transfert radiatif SAIL pour réaliser une assimilation de réflectances en utilisant l'algorithme même du simplexe (Nelder & Mead, 1965). Ils estiment ainsi cinq variables d'entrée et ne prennent pas en compte d'information *a priori*.

L'utilisation d'une information *a priori* présente deux avantages : (i) d'estimer davantage de variables d'entrée, ce qui permet de mieux traduire les interactions sol-plante complexes et (ii) d'éviter que certaines valeurs de variables d'entrée ne prennent des valeurs aberrantes ou ne jouent le rôle de variables d'ajustement : l'information *a priori* doit permettre aux variables de conserver leur sens physique. Les trois méthodes d'assimilation que nous comparerons permettront donc toutes d'intégrer ce genre d'information.

La troisième étape est de simuler les variables d'intérêt final à partir des variables d'entrée estimées par assimilation. Ces variables sont celles dont a été l'objet le Chapitre II : le rendement, la teneur en protéine et le reliquat d'azote à la récolte. Ces simulations seront

comparées aux simulations réalisées à partir de l'IAP : le critère d'évaluation des méthodes d'assimilation sera le degré d'amélioration de la qualité des simulations par assimilation par rapport à l'IAP.

Le plan de la suite de cette partie est le suivant :

- présentation des trois méthodes d'assimilation comparées ;
- comparaison des performances des trois méthodes d'assimilation sur un jeu de données simulées et sélection d'une d'entre elles ;
- pour compléter l'analyse, les effets de la configuration du jeu de données servant à l'assimilation sur les résultats de l'assimilation seront évalués avec la méthode retenue.

III.3.2 Présentation des trois méthodes d'assimilation utilisées

On ne comprend pas la science de la chaussure
quand on ne comprend pas ce qu'est la science
Platon, Théétête

Tous les programmes nécessaires à l'application à STICS des méthodes décrites ont été réalisés avec le logiciel MatLab® (The MathWorks, Inc.).

III.3.2.1 Méthode de minimisation d'un critère bayésien par l'algorithme du simplexe

Avec ce type de méthodes, on cherche à estimer grâce à un algorithme de recherche les valeurs des variables d'entrée du modèle qui minimisent un critère traduisant la différence entre les valeurs des variables observées et simulées par le modèle. Le simplexe a été choisi ici comme algorithme de recherche car il est largement utilisé et adapté à la résolution de problèmes présentant des discontinuités (aide MatLab), ce qui est le cas de STICS (Prévot *et al.*, 2003). La fonction de coût utilisée, ou critère à minimiser, introduit l'information *a priori* selon la méthode décrite par Makowski & Jeuffroy (2002). Cette méthode sera désignée, par souci de commodité, SIMPLEXE.

III.3.2.1.1 Description de l'algorithme du Simplexe

La description qui est faite ici du simplexe est tirée de Sorooshian & Gupta (1993). Cet algorithme a été initialement développé par Nelder & Mead (1965).

Soit J le nombre de variables d'entrée du modèle que l'on veut estimer. Les différentes étapes de l'algorithme décrites ci-dessous sont illustrées par la Figure III.3-1 dans le cas où J=2.

(0) Sélection de J+1 vecteurs θ de variables d'entrée dans l'espace des valeurs possibles et calcul de la fonction à minimiser (fonction coût) pour chacun de ces vecteurs qui forment le simplexe. Les fonctions coût sont représentées par les points X_1, X_2 et X_3 dans l'espace des valeurs des variables d'entrée. Certains modes de réalisation de cet algorithme demandent un point de départ, θ_0, autour duquel sont choisis ces J+1 vecteurs θ. Figure III.3-1-a.

(1) Identification du vecteur qui donne la plus forte valeur de la fonction coût (représenté par le point X_3 dans la Figure III.3-1-a).

(2) Calcul de la fonction coût pour le barycentre des J meilleurs vecteurs θ (représenté par le point X_{ce}, barycentre des points X_1 et X_2, dans la Figure III.3-1-a).

(3) Localisation d'un nouveau vecteur par **réflexion** du vecteur représenté par X_3 par rapport au vecteur représenté par X_{ce} : point X_r (Figure III.3-1-a). Si la valeur de la fonction coût pour le vecteur représenté par X_r est meilleure que celle obtenue avec le vecteur représenté par X_3, aller en (4) ; sinon, aller en (5).

(4) Localisation d'un nouveau vecteur par **expansion** d'un facteur 2 de l'étape de réflexion (Figure III.3-1-b) : point X_e. Si la valeur de la fonction coût pour le vecteur représenté par le point X_e est meilleure que celle obtenue avec le vecteur représenté par X_r, remplacer X_3 par X_e ; sinon, remplacer X_3 par X_r. Aller en (7).

(5) Localisation par **contraction** d'un nouveau vecteur situé à mi-distance des vecteurs associés à X_3 et X_{ce} (Figure III.3-1-c) : point X_c. Si la valeur de la fonction coût pour le vecteur représenté par X_c est meilleure que celle obtenue avec le vecteur représenté par X_3, remplacer X_3 par X_c et aller en (7) ; sinon, aller en (6).

(6) **Réduction** du simplexe en déplaçant chacun des J moins bons vecteurs (donc tous sauf celui correspondant à X_1) à mi-chemin entre leur position actuelle et celle du vecteur représenté par X_1 (Figure III.3-1-d) : points $X_{2'}$ et $X_{3'}$.

(7) Répétition des étapes (1) à (6) jusqu'à ce que la taille du simplexe devienne plus petite que le critère de convergence. A ce moment-là, le vecteur associé à la meilleure valeur de la fonction coût est une estimation de l'optimum.

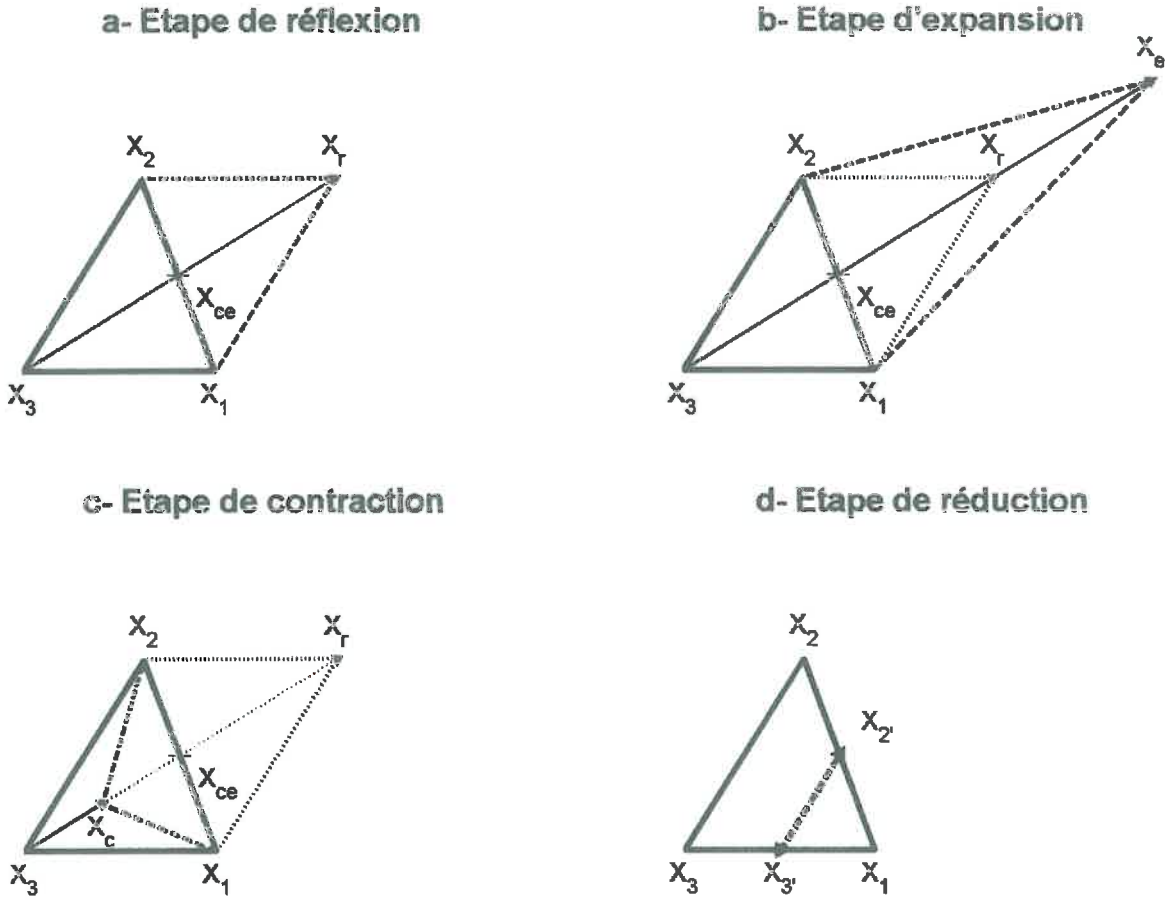


Figure III.3-1 : Algorithme de recherche du simplexe. X_1 : point représentant le vecteur donnant la plus faible valeur de la fonction coût ; X_2 : point représentant le vecteur donnant la seconde plus faible valeur de la fonction coût ; X_3 : point représentant le vecteur donnant la plus forte valeur de la fonction coût ; X_{ce} : barycentre des deux meilleurs points X_1 et X_2 ; X_r : point obtenu par réflexion de X_3 par rapport à X_{ce} ; X_e : point obtenu par expansion d'un facteur 2 de X_r ; X_c : point obtenu par contraction, situé entre X_3 et X_{ce} ; $X_{2'}$ et $X_{3'}$: points obtenus par réduction de X_2 et X_3 respectivement vers X_1. Extrait de Sorooshian & Gupta (1993).

III.3.2.1.2 Le critère bayésien utilisé

Le critère à minimiser que nous avons utilisé a été proposé par Makowski & Jeuffroy (2002). Il s'écrit, dans notre cas, de la façon suivante :

$$F(\theta)=\sum_{d=1}^{D}\left(\frac{LAId-M_{LAI}(\theta,X_d)}{\sigma_{LAI,d}^{vo}}\right)^2+\sum_{d=1}^{D}\left(\frac{QNd-M_{QN}(\theta,X_d)}{\sigma_{i,d}^{vo}}\right)^2+\sum_{j=1}^{J}\left(\frac{\theta_j-\mu_j^{ve}}{\sigma_j^{ve}}\right)^2 \qquad \text{Équation III.3-1}$$

où D désigne le nombre de dates d'observations, LAI_d et QN_d respectivement les valeurs de LAI et de QN observées à la date d, $M_{LAI}(\theta,X_d)$ et $M_{QN}(\theta,X_d)$ respectivement les valeurs

simulées par le modèle de *LAI* et de *QN* à la date *d* pour le jeu de variables d'entrée θ, $\sigma^{vo}_{LAI,d}$ et $\sigma^{vo}_{QN,d}$ les écarts types des variables observées *LAI* et *QN*, *J* le nombre de variables d'entrée estimées (taille du vecteur θ), θ_j la valeur testée par l'algorithme de la variable d'entrée *j*, μ^{ve}_j et σ^{ve}_j respectivement l'espérance et l'écart type de la distribution *a priori* de la variable d'entrée θ_j. Dans cette approche, les valeurs μ^{ve}_j et σ^{ve}_j pour chaque variable d'entrée θ_j constituent l'information *a priori*.

En supposant que la distribution *a priori* des *J* variables d'entrée est définie comme une loi normale $N(\mu^{ve}_j ; \sigma^{ve}_j)$ et que les erreurs du modèle sont indépendantes et distribuées selon une loi normale $N(0 ; \sigma^{vo}_{LAI,d})$ et $N(0 ; \sigma^{vo}_{QN,d})$, respectivement pour *LAI* et *QN*, le mode de la distribution *a posteriori* est obtenu en minimisant cette expression.

Dans le cadre de l'étude que nous nous sommes proposé, l'information *a priori* est constituée d'un intervalle $[a_j ; b_j]$. La moyenne μ^{ve}_j de la distribution normale peut être placée au centre de cet intervalle. Quant à l'écart type, σ^{ve}_j, il peut être calculé en supposant que a_j et b_j correspondent à des quantiles extrêmes de la loi normale, par exemple 5% et 95%. Plus simplement, et ce qui revient à choisir une valeur particulière des quantiles, nous l'avons fixé à la valeur de l'écart type de la loi uniforme. Nous avons donc supposé que :

$$\forall j \in [1;J] \quad \mu^{ve}_j = \frac{a_j + b_j}{2}$$ Équation III.3-2

$$\forall j \in [1;J] \quad \sigma^{ve}_j = \sqrt{\frac{(b_j - a_j)^2}{12}}$$ Équation III.3-3

III.3.2.1.3 Normalisation des variables d'entrée du modèle estimées et vecteurs initiaux

Dans son fonctionnement, l'algorithme d'optimisation requiert des valeurs normalisées des variables d'entrée. La normalisation a été effectuée pour chaque variable d'entrée θ_j par homothétie de la gamme de variation donnée par l'information *a priori* vers la gamme [0;1]. Cette homothétie revient simplement à appliquer la relation suivante : $\theta^{norm}_j = (\theta_j - a_j)/(b_j - a_j)$ dans laquelle θ^{norm}_j est la valeur normalisée de θ_j, et $[a_j;b_j]$ représente la gamme de variation de la distribution *a priori*. Au moment du calcul du critère et de l'appel du modèle, cette variable doit avoir sa valeur naturelle, avant normalisation.

L'algorithme du simplexe tel qu'il est développé dans la fonction *fminsearch* de MatLab nécessite un vecteur initial. Une étude préalable ayant montré l'importance de cette valeur de départ, pour une assimilation, 25 ont été effectuées en réalité, à partir de 25 vecteurs initiaux tirés aléatoirement entre les bornes des valeurs de variables d'entrée $[a_j ; b_j]$. L'assimilation conduisant au vecteur final donnant le plus faible critère a été retenue. Le choix de n'utiliser que 25 vecteurs initiaux est contestable, vu la sensibilité du résultat de l'estimation au vecteur initial.

III.3.2.2 Méthode stochastique et bayésienne : GLUE

III.3.2.2.1 Présentation générale de la méthode

La méthode GLUE (Generalized Likelihood Uncertainty Estimation) a été développée dans le domaine de l'hydrologie pour estimer les variables d'entrée de modèles complexes (Beven & Binley, 1992). Beven & Freer (2001) y voient une méthode permettant de prendre implicitement en compte les effets (i) de non linéarité des modèles, (ii) de covariance des valeurs des variables d'entrée, (iii) des erreurs des données d'entrée ou des variables observées avec lesquelles les simulations sont comparées ou des erreurs dans la structure du modèle. Ces mêmes auteurs soulignent que les choix qui doivent être faits pour mettre en œuvre cette méthode sont limités et doivent être explicitement discutés. Elle est apparentée par Kuczera & Parent (1998) à une approche de type Monte-Carlo connue sous le nom de « importance sampling ». Elle consiste à estimer une distribution *a posteriori* des valeurs des variables d'entrée du modèle que l'on veut estimer en utilisant le théorème de Bayes (*cf.* Chapitre I) d'après l'information *a priori* et les valeurs observées.

III.3.2.2.2 Principe et mise en œuvre de la méthode

La méthode GLUE est ici décrite d'après Makowski *et al.* (2002).

Cette méthode, basée sur le théorème de Bayes, consiste à déterminer une approximation de la distribution *a posteriori*, notée $\pi(\theta|Y)$, des variables d'entrée θ par une distribution discrète (θ_n, p_n), où p_n désigne un poids affecté à une réalisation particulière θ_n de θ. Y désigne l'ensemble des observations.

Autrement dit, N_G valeurs possibles de vecteurs de variables d'entrée θ_n seront affectées d'une probabilité p_n avec $n=1,\ldots,N_G$ et $\sum_{n=1}^{N_G} p_n=1$. Ces N_G valeurs possibles sont tirées aléatoirement dans des distributions qui sont définies par l'information *a priori*. La

connaissance de p_n permet d'avoir des informations sur la distribution *a posteriori* de chaque variable θ_j.

L'algorithme est le suivant :

(1) Tirage aléatoire de N_G vecteurs θ_n, n=1,...,N_G, dans la distribution *a priori*. A partir de ces N_G vecteurs, on réalise N_G simulations donnant les valeurs simulées des I variables observées, notées $M_i(\theta_n, X_d)$.

(2) Calcul des valeurs de vraisemblance $\pi(Y|\theta_n)$ et des densités *a priori* $\pi(\theta_n)$, associées aux différents vecteurs générés, n=1,...,N_G.

- Si les erreurs du modèle sont indépendantes et distribuées selon une loi normale, la fonction de vraisemblance $\pi(Y|\theta_n)$ se calcule de la façon suivante pour chaque vecteur de variables d'entrée θ_n :

$$\pi(Y|\theta_n)=\prod_{d=1}^{D}\frac{1}{\sqrt{2\pi\sigma_{LAI,d}^{vo\,2}}}\exp\frac{(LAI_d-M_{LAI}(\theta_n,X_d))^2}{2\sigma_{LAI,d}^{vo\,2}}\prod_{d=1}^{D}\frac{1}{\sqrt{2\pi\sigma_{QN,d}^{vo\,2}}}\exp\frac{(QN_d-M_{QN}(\theta_n,X_d))^2}{2\sigma_{QN,d}^{vo\,2}}$$

Équation III.3-4

où les notations sont les mêmes que celles utilisées dans la description du critère à minimiser de la méthode SIMPLEXE.

- $\pi(\theta_n)$ se calcule à partir de l'information *a priori*.

(3) Calcul de la probabilité *a posteriori* p_n associée au vecteur θ_n :

$$p_n=\frac{\pi(Y|\theta_n)\pi(\theta_n)}{\sum_{j=1}^{N}\pi(Y|\theta_j)\pi(\theta_j)}$$

Équation III.3-5

d'après le théorème de Bayes.

Les couples (θ_n, p_n), n=1,...,N_G peuvent être utilisés pour calculer par exemple la moyenne $\bar{\theta}$ et l'écart type $\bar{\sigma}$ de la distribution *a posteriori* :

$$\bar{\theta}=\sum_{n=1}^{NG}p_n\cdot\theta_n$$

Équation III.3-6

$$\overline{\sigma}^2 = \sum_{n=1}^{N_G} p_n \cdot \theta_n^2 - \left(\sum_{n=1}^{N} p_n \cdot \theta_n \right)^2$$ Équation III.3-7

L'utilisation de cette méthode ne requiert que la définition (i) du nombre de jeux de variables d'entrée à générer et de simulations à réaliser, N_G, (ii) des distributions *a priori* des différentes variables d'entrée et (iii) des valeurs $\sigma_{LAI,d}^{vo}$ et $\sigma_{QN,d}^{vo}$.

III.3.2.2.3 Utilisations antérieures de la méthode

Cette méthode, d'après Makowski *et al.* (2002), a essentiellement été utilisée en hydrologie. Kuczera & Parent (1998) la comparent à une autre méthode de type Monte-Carlo, l'algorithme de Metropolis, sur un modèle hydrologique. Ils trouvent cette dernière meilleure et soulignent que la méthode GLUE peut avoir l'inconvénient de nécessiter un grand nombre de simulations pour correctement caractériser l'espace des variables d'entrée à estimer si celles-ci sont nombreuses. Makowski *et al.* (2002) comparent la méthode GLUE à l'algorithme de Metropolis-Hastings sur un modèle de culture simplifié. Ils trouvent que ces deux méthodes ont des qualités semblables et que leurs résultats sont très dépendants de la valeur donnée aux informations *a priori*. Cependant, des développements plus récents montrent que la méthode GLUE serait moins performante que Metropolis-Hastings (Makowski, comm. pers.).

L'hypothèse d'indépendance et de distribution normale des erreurs du modèle est forte dans le cas d'un modèle de culture : les erreurs faites sur *LAI* et *QN* d'une date à l'autre et entre ces deux variables sont très probablement liées. Ce fait sera à prendre en compte dans l'interprétation des résultats. Néanmoins, on pourrait prendre en compte la corrélation entre les erreurs en cherchant à les estimer au même titre que les variables d'entrée.

Un avantage essentiel de cette méthode est que le nombre de simulations à réaliser lors de l'étape (1) de l'algorithme, N_G, qui peut être très important comme le soulignent Kuczera & Parent (1998), ne dépend pas du nombre de points observés. Autrement dit, on peut utiliser les mêmes N_G simulations pour réaliser l'assimilation concernant tous les pixels d'une même parcelle pour lesquels l'information *a priori* est la même, alors que pour la plupart des autres méthodes, l'algorithme de recherche doit être appliqué pixel par pixel, car au cours de sa progression, les simulations à réaliser sont influencées par les observations (par exemple dans l'algorithme de Metropolis-Hastings ou du simplexe). De ce point de vue, c'est donc une méthode particulièrement bien adaptée à la problématique de l'agriculture de précision. Le

calcul des vraisemblances à l'étape (2), qui doit être réalisé pour tous les pixels avec les N_G simulations, est quant à lui extrêmement rapide (dans notre application, dix minutes suffisent à ce calcul contre un jour et demi pour réaliser $N_G = 200\ 000$ simulations).

III.3.2.3 Méthode stochastique basée sur une analyse de sensibilité : MCIP

Récoltez d'abord l'ivraie et liez-la en bottes pour la consumer ;
quant au blé, ramassez-le dans mon grenier.
Matthieu, 13, 30.

La présentation de la méthode MCIP (Multi-objective Calibration Iterative Procedure), développée par Demarty (2001), se base essentiellement sur ce qui est rapporté par Demarty *et al.* (2004a et b), car c'est d'après ce travail que cette méthode a été adaptée à notre problème.

III.3.2.3.1 Présentation générale de la méthode

Les bases théoriques initiales de cette méthode sont données par Yapo *et al.* (1998) et Gupta *et al.* (1998).

Deux idées essentielles forment son fondement :

- elle utilise une approche multicritère de l'estimation des paramètres : elle cherche à minimiser conjointement plusieurs critères, sans les regrouper dans un critère unique par pondération. Comme la minimisation simultanée de plusieurs critères ne permet pas d'isoler une solution unique, on cherche à déterminer un ensemble de solutions équivalentes du point de vue multicritère, au sens où elles apportent des améliorations sur certains d'entre eux tout en en dégradant d'autres ;

- elle réalise un choix parmi les variables d'entrée estimées en déterminant celles dont les valeurs permettent de réaliser des simulations en bon accord avec un ensemble d'observations. C'est en ce sens qu'on doit entendre ici le terme d'analyse de sensibilité qui sera discuté plus tard.

La méthode MCIP est dérivée d'une approche d'analyse de sensibilité multicritère, MOGSA (MultiObjective Generalized Sensitivity Analysis) développée par Bastidas *et al.* (1999).

III.3.2.3.2 Principe et mise en œuvre de la méthode

L'algorithme de cette méthode peut être décrit en trois étapes :

(1) Création d'un jeu de N_M vecteurs θ_n $n = 1,\ldots,N_M$, de variables d'entrées qui sera scindé en deux sous-ensembles : (i) un sous-ensemble d'effectif N_a composé de vecteurs conduisant à des simulations « acceptables », c'est-à-dire qui minimisent au sens multicritère les différences entre observations et simulations ; (ii) un sous-ensemble d'effectif N_{na} composé de vecteurs conduisant à des simulations « non acceptables ».

(2) Sélection, au sein de l'ensemble des variables d'entrée, de celles auxquelles la qualité des simulations du modèle est sensible en se basant sur les deux sous-ensembles déterminés à l'étape (1) et réduction de leurs intervalles de variation initiaux vers des intervalles qui favorisent la qualité des simulations.

(3) Itération des étapes (1) et (2) jusqu'à ce que les simulations soient suffisamment proches des observations.

III.3.2.3.2.1 Création de deux sous-ensembles de valeurs de variables d'entrée conduisant à des simulations « acceptables » et « non-acceptables »

N_M simulations sont réalisées de façon similaires à la méthode GLUE : N_M vecteurs de variables d'entrée θ_n ($n = 1,\ldots,N_M$) sont associés à N_M vecteurs composés des variables simulées $M_{LAI,d}(\theta_n,X_d)$ et $M_{QN,d}(\theta_n,X_d)$ correspondant aux observations LAI_d et QN_d pour les différentes dates d'observation d. Les valeurs des J variables d'entrée sont échantillonnées selon une loi uniforme dans $[a_j\,;\,b_j]$, $j = 1,\ldots,J$.

La partition des N_M simulations en deux sous-ensembles A={Simulations « acceptables »} et NA={Simulations « non-acceptables »}, d'effectifs N_a et N_{na} respectivement, se fait grâce à une approche multicritère. Une telle approche cherche à minimiser conjointement plusieurs fonctions de θ qui mesurent les écarts entre les simulations et les observations. Dans notre étude, par souci d'homogénéité avec les deux autres méthodes présentées ci-dessus, nous avons choisi de minimiser des RMSE pondérées par les valeurs $\sigma^{vo}_{LAI,d}$ et $\sigma^{vo}_{QN,d}$:

$$\underset{\theta}{Min}\left\{\sum_{d=1}^{D}\left(\frac{LAId-M_{LAI}(\theta,X_d)}{\sigma^{vo}_{LAI,d}}\right)^2,\sum_{d=1}^{D}\left(\frac{QNd-M_{QN}(\theta,X_d)}{\sigma^{vo}_{QN,d}}\right)^2\right\} \qquad \text{Équation III.3-8}$$

où les notations sont les mêmes que celles employées pour les méthodes SIMPLEXE et GLUE.

En général, une simple RMSE est choisie pour exprimer ces fonctions à minimiser conjointement (Demarty *et al.*, 2004a ; Yapo *et al.*, 1998). D'autres critères statistiques pourraient être utilisés pour une même variable : on pourrait par exemple chercher à minimiser conjointement la RMSE et l'efficience des différentes variables observées (Yapo *et al.*, 1998). Face à la diversité des observations disponibles et des critères statistiques permettant d'évaluer un modèle, cette approche peut être intéressante car elle permet d'optimiser plusieurs critères simultanément pour une ou plusieurs variables observées.

Comme il est peu probable d'arriver à déterminer une valeur de θ qui minimise à la fois toutes les fonctions coûts, on cherche à déterminer un ensemble de solutions qui apportent alternativement une amélioration sur un critère et une détérioration sur un autre. Cette façon de faire est préférée à celle qui consisterait à créer un critère synthétique par assignation de poids aux différentes fonctions à minimiser, ce qui ne pourrait être que subjectif (Madsen, 2000). Cette approche correspond au rangement de Pareto (Gupta *et al.*, 1998).

Chaque simulation est représentée dans l'espace défini par les fonctions à minimiser (Equation IV.3-8). Un exemple conceptuel est donné dans la Figure III.3-2-a. Afin de séparer les simulations « acceptables » des simulations « non-acceptables », on procède à un rangement de Pareto, défini de la façon suivante par Gupta *et al.* (1998) :

- pour chaque non-membre de l'ensemble de Pareto, il existe au moins un membre de l'ensemble de Pareto dont tous les critères sont meilleurs ;

- pour chaque membre de l'ensemble de Pareto, il est impossible de trouver un autre membre de cet ensemble dont tous les critères sont meilleurs.

Sur l'exemple de la Figure III.3-2-a, les points encerclés font nécessairement partie de l'ensemble de Pareto car ils minimisent l'un des deux critères. Si l'on considère par exemple celui qui minimise la RMSE pondérée du *LAI*, le point suivant faisant partie de l'ensemble de Pareto améliore la RMSE pondérée sur *QN* tout en dégradant la première. Les différences entre les niveaux d'amélioration et de dégradation ne sont pas pris en compte, car ce serait introduire implicitement une pondération entre les variables. Aussi ne peut-on pas dire lequel

de ces deux points est meilleur que l'autre : ils font donc partie du même ensemble de Pareto. De même, les deux points encadrés d'un carré font partie de ce même ensemble ; celui qui est mis en exergue par un losange n'en fait quant à lui pas partie car, s'il est vrai qu'il améliore le RMSE de *QN* et et dégrade la RMSE de *LAI* par rapport au carré plein, il dégrade les deux critères par rapport au carré vide. L'ensemble de Pareto de rang 1 est ainsi déterminé.

En mettant de côté cet ensemble, il devient possible de reproduire la procédure avec les points restants : on détermine ainsi les rangs de Pareto de rang 2 et plus, jusqu'à ce que tous les points aient été rangés (Figure III.3-2-b). Ensuite, les simulations peuvent être scindées en deux sous-ensembles, A={Simulations « acceptables »} et NA={Simulations « non-acceptables »} en choisissant un rang de Pareto seuil. Dans la figure exemple, le rang 2 a été sélectionné pour réaliser cette séparation, toutes les simulations appartenant aux ensembles de Pareto de rangs 1 ou 2 sont affectées au sous-ensemble A, les autres à B.

Afin d'éviter d'avoir affaire à de trop fortes valeurs de fonctions objectifs, on peut également choisir d'appliquer un seuil (Leplastrier *et al.*, 2002) : dans l'exemple montré, on pourrait choisir d'éliminer de l'ensemble des simulations « acceptables » toutes celles qui conduisent à une valeur de RMSE pondérée de *LAI* supérieure à 0.6 et de RMSE pondérée de *QN* supérieure à 30.

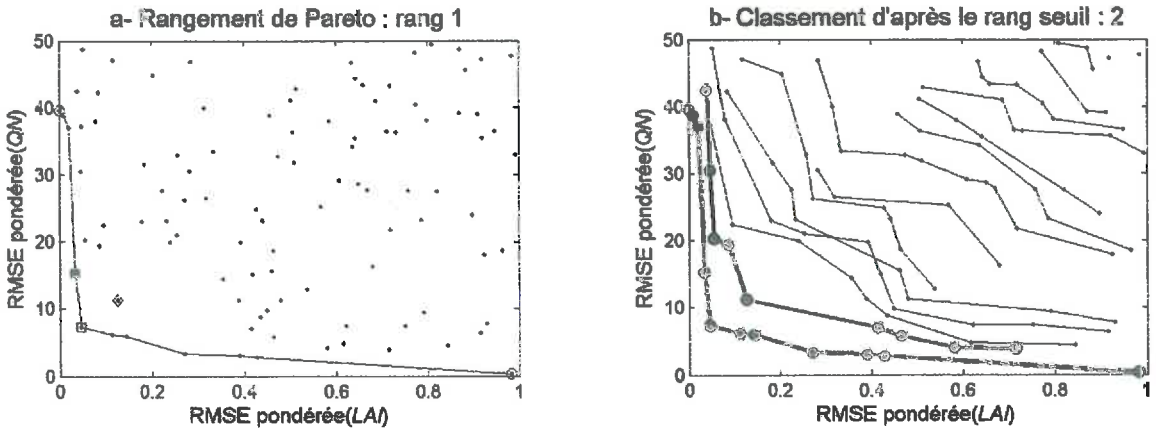


Figure III.3-2 : Représentation d'un cas théorique de la mise en œuvre de la méthode MCIP. a- Les différents jeux de variables d'entrée tirés au hasard dans la distribution *a priori* sont représentés dans l'espace des deux fonctions objectifs : la RMSE pondérée calculée pour le *LAI* et celle calculée pour *QN*. L'ensemble de Pareto de rang 1 est matérialisé. Sa détermination commence par les points encerclés. Les points entourés d'un carré et d'un losange sont commentés dans le texte ; b- Tous les ensembles de Pareto sont matérialisés, du rang 1 au rang 15. Le rang seuil d'ordre 2 est choisi pour scinder l'ensemble des simulations en un sous-ensemble de simulations « acceptables » (traits épais) et « non-acceptables » (traits fins).

III.3.2.3.2.2 Sélection des variables d'entrée auxquelles la qualité des simulations est sensible et réduction de leur intervalle de variation

Une fois l'ensemble des simulations scindé en deux sous-ensembles, A et NA, d'effectifs respectifs N_a et N_{na}, on procède à l'analyse, pour chaque variable d'entrée θ_j de θ, de la répartition de ses différentes valeurs au sein des deux sous-ensembles. Dans l'exemple abstrait proposé dans la Figure III.3-3, le graphique a- montre pour une variable d'entrée donnée, θ_j, les fonctions de distribution dans les deux sous-ensembles A et NA des valeurs prises par la variable.

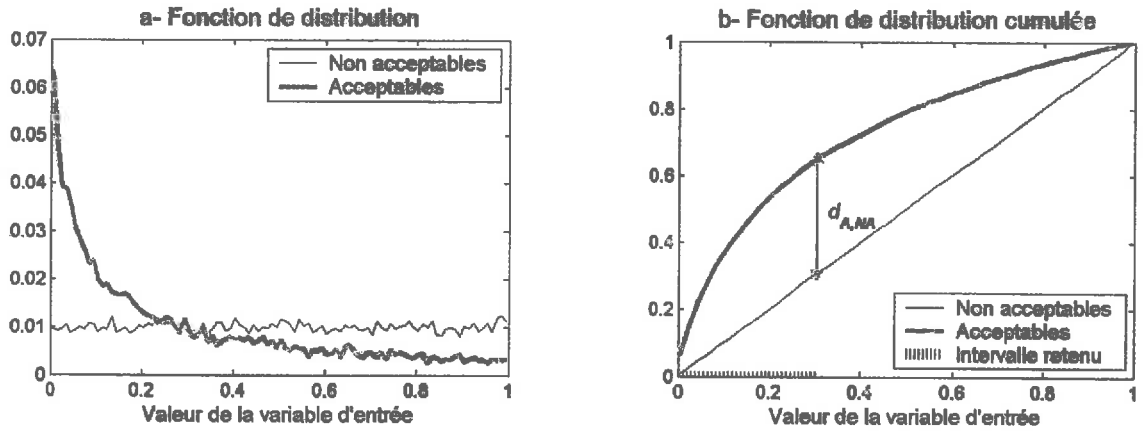


Figure III.3-3 : Représentation sur un cas théorique de la méthode de réduction de l'intervalle de variation initial d'une variable d'entrée donnée. a- Représentation de la fonction de distribution des valeurs de la variable d'entrée pour les simulations acceptables et non acceptables. L'intervalle de variation initial est de [0;1]. Les simulations acceptables sont données par des valeurs plutôt faibles de la variable ; b- Représentation de la distribution cumulée des valeurs de la variable d'entrée pour les simulations acceptables et non acceptables. La distance $d_{Na,Nna}$ permet de mesurer la sensibilité de la variable : plus la fonction de distribution cumulée des simulations acceptables s'éloigne de celle des simulations non acceptables, plus la variable considérée a un poids important sur la qualité des simulations. L'intervalle de variation retenu est celui qui contient les valeurs les plus fréquemment rencontrées dans les simulations acceptables.

On voit que les faibles valeurs de la variable considérée sont plus fortement représentées dans le sous-ensemble A que dans le sous-ensemble NA : non seulement cette variable est sensible dans le sens où ses valeurs influencent la qualité des simulations, mais, en outre, on apprend que ses « meilleures valeurs » sont les plus faibles. Ceci est représenté de façon différente par le graphique b- de la même figure où sont montrées les fonctions de distribution cumulées et normalisées pour une variable donnée θ_j. En trait épais, la fonction de distribution cumulée des simulations acceptables, $P_A(\theta_j)$, se détache de celle, matérialisée par un trait fin, des simulations non acceptables, $P_{NA}(\theta_j)$. Ceci traduit ce qui vient d'être dit : les faibles valeurs de la variable considérée sont plus largement représentées dans le premier

sous-ensemble. Inversement, la qualité des simulations sera jugée peu sensible à une variable d'entrée si, quelle que soit sa valeur, les fonctions de distribution sont proches.

Cette sensibilité du modèle à une variable donnée peut être évaluée par le test statistique de Kolmogorov-Smirnov. La valeur $d_{A,NA}$, qui traduit la distance maximale entre ces deux fonctions cumulées, est d'abord calculée, pour toutes les variables d'entrée θ_j :

$$d_{A,NA}(\theta_j)=\sup_{\theta_j}(P_A(\theta_j)-P_{NA}(\theta_j)) \qquad \text{Équation III.3-9}$$

Elle est représentée sur le graphique par la double flèche. La probabilité de Kolmogorov-Smirnov affectée à chaque variable et qui traduit le niveau de sensibilité du modèle à cette variable peut être approchée de la façon suivante :

$$P_{KS}(\theta_j)=Q_{KS}\left[d_{A,NA}(\theta_j)\cdot\left(\sqrt{N_e}+0.12+\frac{0.11}{\sqrt{N_e}}\right)\right] \qquad \text{Équation III.3-10}$$

avec :

$N_e=\frac{N_a\cdot N_{na}}{N_a+N_{na}}$, nombre effectif Équation III.3-11

$Q_{KS}(x)=2\cdot\sum_{k=1}^{\infty}(-1)^{k-1}\cdot e^{-2k^2x^2}$, fonction de x Équation III.3-12

Les valeurs de P_{KS} obtenues sont comprises entre 0 et 1 ; plus la valeur est faible, plus la qualité des simulations du modèle est jugée sensible à la variable pour laquelle elle a été estimée. A ce stade, il faut choisir un seuil de sensibilité : en-dessus de cette valeur, la qualité des simulations ne sera pas jugée sensible à la variable considérée ; en-dessous, elle le sera et la gamme de variation de la variable sera restreinte comme indiqué sur la Figure III.3-3-b : la nouvelle gamme de variation sera délimitée par la valeur de θ_j qui maximise $d_{Na,Nna}$, notée θ_j^d. On passe ainsi de l'intervalle $[a_j;b_j]$ à l'intervalle $[a'_j;b'_j]$ avec $a'_j=a_j$ et $b'_j=\theta_j^d$ si les meilleures valeurs de la variable θ_j sont les plus faibles comme dans la figure exemple ou $a'_j=\theta_j^d$ et $b'_j=b_j$ dans le cas inverse.

La notion d'« analyse de sensibilité » développée ici se démarque de sa définition habituelle. Alors que classiquement les analyses de sensibilité s'intéressent aux effets des valeurs des variables d'entrée du modèle sur ses sorties indépendamment de toute observation, et ce pour l'ensemble des gammes de variation des variables d'entrée, cette

méthode cherche à réduire les intervalles initiaux des variables d'entrée de façon à rapprocher les sorties du modèle des observations, s'apparentant de ce fait à de l'estimation. Elle détermine ainsi quelles sont les variables d'entrée qu'il est important d'estimer en fonction de leur importance sur la qualité des simulations. Il faut donc parler en ce sens de « *sensibilité de la qualité des simulations à une variable d'entrée donnée* », dans un contexte d'observation donné. Toutefois, pour gagner en concision, nous nous permettrons occasionnellement l'abus de langage « *variable sensible* » qui se voudra en être l'équivalent.

Par ailleurs, ces « analyses de sensibilité » peuvent également se faire en mono-critère : un seul critère est alors utilisé pour répartir les simulations effectuées en simulations « acceptables » et « non acceptables » et pour calculer la distance $d_{Na,Nna}$. Cette potentialité ne sera pas exploitée ici, mais certaines figures représentent les résultats d'analyses mono-critères (Figure III.3-5 à Figure III.3-7).

III.3.2.3.2.3 Itération de la méthode et automatisation

Afin d'estimer les variables d'entrée du modèle, Demarty (2001) propose de répéter les deux points précédents, (i) création d'un jeu de données et partition entre simulations « acceptables » et « non-acceptables » et (ii) sélection des variables sensibles et restriction de leurs gammes respectives de variation, afin de contraindre peu à peu les simulations. Le jeu de données créé au début de chaque nouvelle itération le sera avec les nouveaux intervalles de variation. Au cours des itérations, des variables auxquelles la qualité des simulations n'avaient pas été jugées sensibles lors des étapes précédentes peuvent le devenir au fur et à mesure de l'augmentation de la contrainte effectuée. Peu à peu, les gammes de variation d'un nombre croissant de variables seront ainsi réduites.

Le processus itératif est schématisé par la Figure III.3-4 : par analogie avec les méthodes bayésiennes, on peut considérer que l'on part d'une distribution *a priori* des variables d'entrée pour aboutir à une distribution *a posteriori*, qui prend en compte les observations.

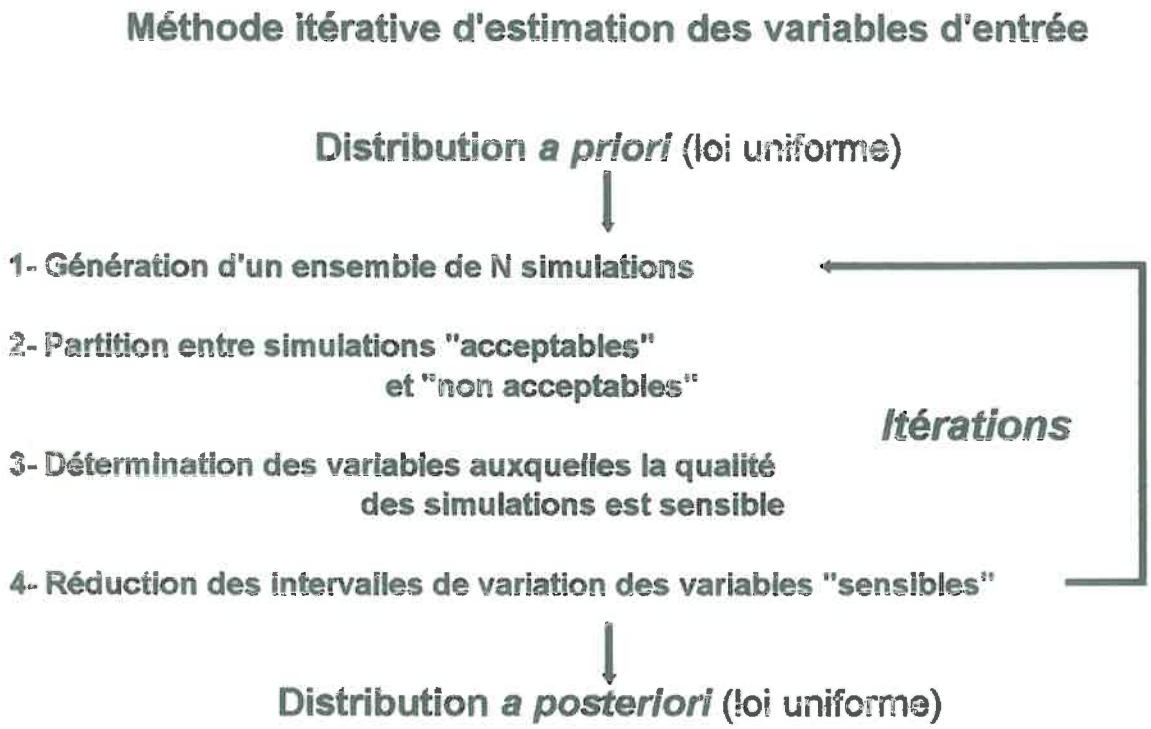


Figure III.3-4 : Principe de l'estimation des variables d'entrée du modèle par applications successives de la méthode MCIP. D'après Demarty (2001).

Les algorithmes permettant de mettre en œuvre cette méthode ont été fournis par J. Demarty. Un travail d'automatisation a été effectué, permettant d'enchaîner les différentes itérations pour plusieurs points traités simultanément, comme l'utilisation dans le contexte de l'agriculture de précision le rendait nécessaire. L'inconvénient de l'automatisation est cependant de supprimer une étape d'analyse par l'utilisateur, instructive sur le fonctionnement du modèle et l'évolution de la gamme de variation des variables d'entrée. Mais pour effectuer une assimilation en pratique dans le cadre de l'agriculture de précision, elle s'avère indispensable.

III.3.2.3.2.4 Synthèse des choix à effectuer lors de l'utilisation de cette méthode

Cette méthode a été mise au point pour éviter d'avoir à fixer des poids entre les différentes fonctions objectifs à pondérer. Néanmoins, comme on a pu le voir, plusieurs choix doivent être faits pour la mettre en œuvre. En voici la synthèse de ceux qui ont été faits dans notre étude :

- distribution *a priori* des variables d'entrée des paramètres : pour toutes les variables d'entrée θ_j, la distribution *a priori* est une loi uniforme définie par les bornes $[a_j ; b_j]$ qui sont supposées connues ;

- nombre N_M de simulations à effectuer : une étude présentée dans la suite a permis de fixer ce nombre à 10 000 ;

- nature des fonctions objectifs (ce choix constitue un des intérêts de la méthode Yapo *et al.*, 1998) : par souci de cohérence avec les deux autres méthodes, nous avons choisi des RMSE pondérées par les erreurs des observations ;

- rang de Pareto qui détermine la limite entre simulations « acceptables » et « non-acceptables » ;

- seuils à fixer aux différentes fonctions objectifs ;

Ces deux derniers points ont fait l'objet de la création d'une procédure permettant de les choisir automatiquement. Dans cette procédure, les seuils des fonctions objectifs sont fixés initialement à la valeur du premier quartile des différentes réalisations de la fonction objectif. Si le nombre des simulations ainsi retenues par seuillage est inférieur à 500, les seuils sont progressivement augmentés jusqu'à retenir au moins 500 simulations. Le rangement de Pareto est ensuite effectué sur les simulations conservées après ce seuillage. Le rang de Pareto fixant la limite entre simulations « acceptables » et « non acceptables » est le plus petit rang tel que les simulations « acceptables » soient au moins au nombre de 200.

- seuil de la valeur de P_{KS} en-dessous duquel la qualité des simulations sera jugée sensible à la variable considérée : il a été fixé à 0.05, conformément à Demarty *et al.* (2004b) ;

- nombre d'itérations à effectuer : six phases ont été réalisées.

Ces choix sont subjectifs pour la plupart d'entre eux mais leur automatisation permet de formaliser la façon de les faire. Ils permettent en outre une certaine souplesse d'utilisation, notamment par le choix des fonctions objectifs.

III.3.2.3.3 Utilisation des méthodes auxquelles la méthode MCIP s'apparente

La méthode telle qu'elle vient d'être présentée est un travail original développé par Demarty (2001). Les bases qu'elle utilise ne correspondent donc pas strictement à celle-ci. Cet auteur l'a appliquée à l'analyse de sensibilité et à l'estimation de variables d'entrée d'un modèle de simulation des transferts sol-végétation-atmosphère. Dans ce travail, les différents

paramètres de la méthode récapitulés ci-dessus sont discutés et la nécessité d'adapter ces différents choix au modèle et à la problématique considérés est mise en évidence.

Bastidas *et al.* (1999) utilisent une méthode proche, mais uniquement pour réaliser une analyse de sensibilité. Ils ne réalisent donc pas d'itération comme montré dans la Figure III.3-4. En revanche, ils utilisent un algorithme qui permet de tester la valeur des différents seuils et paramètres de la méthode : le nombre de simulations à effectuer varie selon les données, les valeurs de 750 et 3000 étant indiquées par ces auteurs ; les fonctions objectifs sont des RMSE des variables observées ; le rang de Pareto qui délimite les deux ensembles de simulations est le $10^{ème}$; ils ne fixent pas de seuil sur les fonctions objectifs ; le seuil de P_{KS} qui juge de la sensibilité de la variable est de 0.01.

Plusieurs auteurs utilisent une méthode apparentée, MOCOM-UA, qui cherche à déterminer l'ensemble de Pareto de rang 1 de la façon la plus complète possible (Yapo *et al.*, 1998 ; Gupta *et al.* 1998 ; Madsen, 2000 ; Madsen *et al.*, 2002). Une fois les simulations de départ effectuées et l'ensemble de Pareto de rang 1 déterminé, toutes les simulations qui n'en font pas partie subissent une optimisation de type simplexe de façon à les faire converger vers l'ensemble de Pareto. Brièvement, cette méthode permet d'obtenir les valeurs de paramètres pour un grand nombre de simulations faisant partie de cet ensemble. Gupta *et al.* (1998) parlent d'environ 25 000 simulations nécessaires avant d'arriver à converger vers cet ensemble de Pareto ; l'étude de Madsen (2000), pour un autre modèle, n'en nécessite que 2000. En ce qui concerne les fonctions objectifs, le critère le plus couramment utilisé est la RMSE (Gupta *et al.*, 1998), qui peut être pondérée (Madsen, 2000). Yapo *et al.* (1998) proposent d'utiliser une RMSE pondérée par des coefficients fonction de la valeur simulée.

III.3.2.4 Conclusion

En conclusion de cette présentation, nous comparerons la façon dont est utilisée l'information que l'on possède *a priori* sur la valeur des variables d'entrée dans les trois méthodes, et la nature de l'information obtenue en sortie sur la distribution *a posteriori* des variables d'entrée.

III.3.2.4.1 Utilisation de l'information *a priori* dans les trois méthodes

On rappelle que dans le cadre que nous avons choisi, l'information *a priori* est composée, pour chaque variable d'entrée θ_j que l'on veut estimer, d'un intervalle : $[a_j;b_j]$.

- Méthode SIMPLEXE : pour toute variable d'entrée j=1,...,J, on détermine à partir de $[a_j;b_j]$ une espérance μ_j^{ve} et un écart type σ_j^{ve} d'une loi normale. μ_j^{ve} et σ_j^{ve} font partie de l'écriture du critère à minimiser et empêchent l'algorithme de s'éloigner trop de μ_j^{ve}, et ce d'autant plus que σ_j^{ve} a une faible valeur.

- Méthode GLUE : on détermine N_G vecteurs de variables d'entrée $\theta_n=\{\theta_n^j\}$, avec j=1,...,J où J représente le nombre de variables d'entrée à estimer, en tirant aléatoirement les θ_n^j selon une loi définie par $[a_j;b_j]$. Ce peut être une loi uniforme par exemple.

- Méthode MCIP : l'information *a priori* est utilisée de la même façon que pour la méthode GLUE.

Il existe donc une différence de nature dans l'utilisation de l'IAP par les méthodes stochastiques (MCIP et GLUE) et la méthode SIMPLEXE. Les premières s'en servent uniquement pour définir l'espace des valeurs possibles qu'elles vont explorer. La méthode SIMPLEXE s'en sert à deux titres : pour définir les valeurs des variables d'entrée autour desquelles elle va effectuer sa recherche (μ_j^{ve}) et pour pondérer les différences (θ_j - μ_j^{ve}) entre elles (σ_j^{ve}).

III.3.2.4.2 Connaissance *a posteriori* des valeurs des variables d'entrée

Chacune des méthodes d'assimilation donne des informations sur la distribution *a posteriori* de chaque variable d'entrée. Mais la nature de ces informations est différente selon les méthodes :

- Méthode SIMPLEXE : pour toute variable d'entrée θ_j, j=1,...,J, elle détermine le mode de sa distribution *a posteriori*, sous l'hypothèse que la distribution *a priori* est normale.

- Méthode GLUE : elle détermine un poids pour chacun des θ_n vecteurs de variables d'entrée, n=1,...,N_G, qui permettra d'estimer, la moyenne et l'écart type de chaque variable d'entrée θ_j, j=1,...,J considérée.

- Méthode MCIP : elle aboutit, pour chaque variable d'entrée j=1,...,J considérée, à un intervalle de distribution final, $[a'_j;b'_j]$, en général réduit par rapport à $[a_j;b_j]$.

Dans les trois méthodes, les valeurs $\sigma^{vo}_{LAI,d}$ et $\sigma^{vo}_{QN,d}$ servent à pondérer les différences entre les valeurs simulées et observées de *LAI* et de *QN* correspondantes.

Ces trois méthodes, qui présentent des différences plus ou moins grandes mais dont le point commun est d'estimer les variables d'entrée du modèle, vont être comparées sur leur faculté à retrouver des valeurs de variables d'entrée ayant servi à réaliser des simulations de référence.

III.3.3 Comparaison des trois méthodes sur un jeu de données simulées

Parmi toutes les variables d'entrée caractéristiques du sol que nous avons définies, toutes n'ont pas la même importance. Nous avons vu que la méthode MCIP était capable de sélectionner les variables qu'il était le plus important d'estimer. Nous verrons que la méthode GLUE agit de façon très analogue, mais ce n'est pas le cas de la méthode SIMPLEXE. Il était donc nécessaire de réaliser une analyse de sensibilité afin de sélectionner un nombre réduit de variables d'entrée à estimer avant de procéder à la comparaison proprement dite. La méthode MCIP a été utilisée pour ce faire. Une fois les variables d'entrée choisies, nous présenterons le protocole de comparaison établi et ses résultats.

III.3.3.1 Sélection des variables d'entrée estimées par assimilation

III.3.3.1.1 Méthodologie et mise en œuvre de la méthode MCIP

Afin de réaliser l'analyse de sensibilité permettant de sélectionner parmi les variables d'entrée listées ci-dessus celles qu'il sera le plus pertinent d'estimer par assimilation, il est nécessaire de créer un jeu de données simulées. Celui-ci est constitué d'un vecteur, θ^{ref}, contenant les 29 variables d'entrée dont on veut étudier la sensibilité et utilisé comme référence, et d'un vecteur contenant les valeurs des variables *LAI* et *QN* simulées à partir de θ^{ref} à 15 dates situées au cours de la partie la plus importante du cycle cultural : 16/02/00, 01/03/00, 15/03/00, 29/03/00, 05/04/00, 12/04/00, 19/04/00, 26/04/00, 03/05/00, 10/05/00, 17/05/00, 24/05/00, 07/06/00, 21/06/00, 05/07/00. Leur fréquence est plus importante aux mois d'avril et mai (7 jours au lieu de 14). Ces valeurs jouent le rôle des observations et sont nommées pour cette raison « pseudo-observées ». Un vecteur composé des variables de sortie d'intérêt final (rendement, teneur en protéines et reliquat), est également associé à θ^{ref}. Il permettra de juger de l'impact de la sélection des simulations par la méthode MCIP sur ces

variables de sortie d'intérêt pratique. Les valeurs contenues par θ^{ref} ont été choisies dans le jeu de données déterminé par l'approche cartographique.

Au cours de cette étude, nous vérifierons la stabilité de la sensibilité de la qualité des simulations aux variables d'entrée en fonction de plusieurs facteurs :

- le nombre de simulations initiales, N_M, réalisées dans la méthode MCIP ; dans l'analyse de base, on aura $N_M = 10\,000$; on étudiera les résultats d'analyse de sensibilité effectuées avec $N_M = 20\,000$; 50 000 ; 100 000 ;

- le nombre d'itérations pratiquées dans la méthode MCIP ; dans l'analyse de base, une seule itération sera utilisée ; on étudiera les résultats obtenus en enchaînant six itérations ;

- la qualité de l'information *a priori* (IAP) ; dans l'analyse de base, on utilisera pour $[a_j ; b_j]$ des intervalles « larges », correspondant à une information disponible classiquement sur une parcelle ; on étudiera les résultats obtenus avec des intervalles « étroits », encadrant de ±10% la valeur de référence ;

- l'année climatique utilisée pour réaliser les assimilations ; dans l'analyse de base, l'année climatique 1999-2000, faisant partie du jeu de données de notre étude, est utilisée. Cette année est particulièrement humide ; on étudiera les résultats obtenus avec une autre année climatique, 1975-1976, marquée quant à elle par une importante sécheresse.

L'effet de la valeur du jeu de données utilisé pour effectuer la simulation de référence, θ^{ref}, sur les résultats de l'analyse de sensibilité n'a en revanche pas été testé.

III.3.3.1.2 Influence du nombre de simulations initiales sur l'analyse de sensibilité

Plusieurs analyses ont été menées utilisant 10 000, 20 000, 50 000 ou 100 000 simulations initiales. Pour l'étude sur 10 000 simulations, trois séries ont été utilisées afin d'évaluer l'impact de l'échantillonnage dans l'espace des valeurs possibles des variables d'entrée. Les résultats de la comparaison sont indiqués dans le Tableau III.3-2.

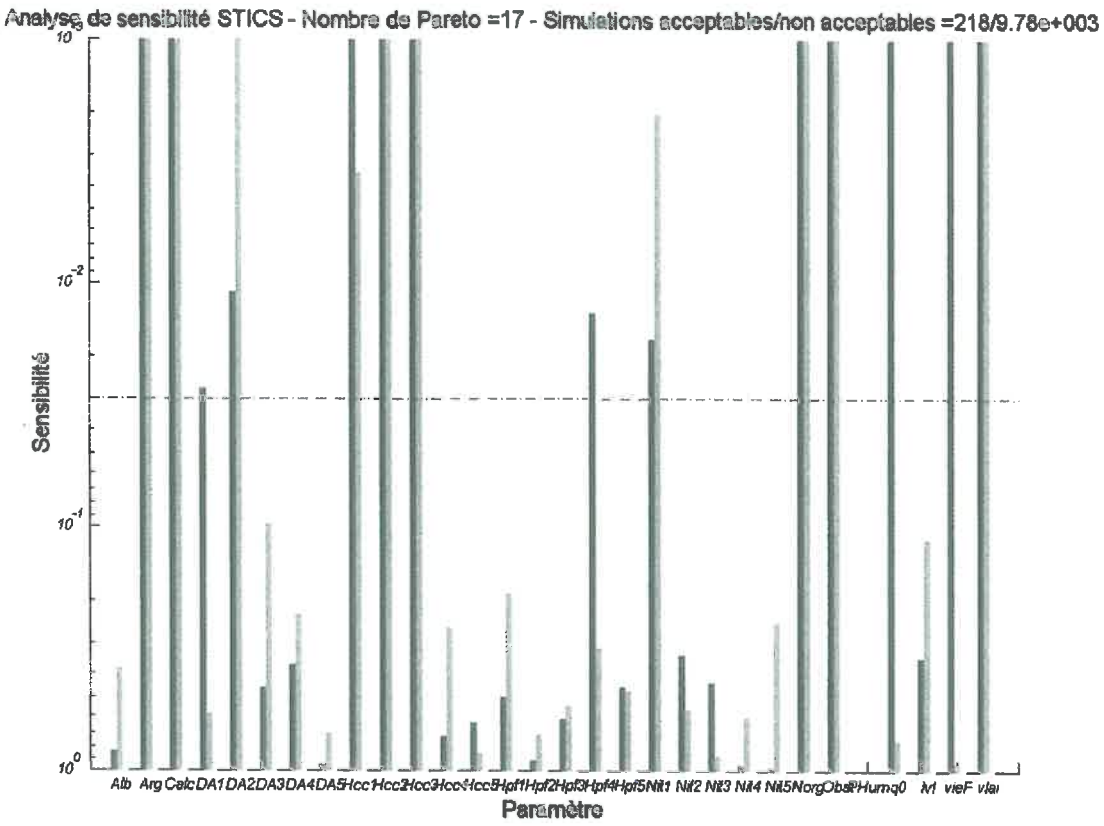


Figure III.3-5 : Exemple d'analyse de sensibilité par la méthode MCIP pour 10 000 simulations de base. Année climatique utilisée : 1999-2000 ; information *a priori* : intervalles larges ; nombre d'itération : une. Les noms des variables d'entrée présentés en abscisse sont respectivement : *ALBEDO, ARGI, CALC, DA(1), DA(2), DA(3), DA(4), DA(5), HCCF(1), HCCF(2), HCCF(3), HCCF(4), HCCF(5), HMINF(1), HMINF(2), HMINF(3), HMINF(4), HMINF(5), NO3initf(1), NO3initf(2), NO3initf(3), NO3initf(4), NO3initf(5), NORG, OBSTARAC, PROFHUM (neutralisée), q0, lvfront, durvieF* et $vlai_{max}$. Les barres noires représentent les sensibilités multicritères (seules prises en compte dans notre étude) et les grises, mono-critères.

Le degré de sensibilité de la qualité des simulations aux variables d'entrée est traduit par des astérisques : *** : forte sensibilité (P_{KS}<0.001) ; ** : sensibilité moyenne (0.001<P_{KS}<0.05) ; * : sensibilité faible (P_{KS}>0.05). La Figure III.3-5 donne un exemple de sortie graphique de la sensibilité dans le cas de 10 000 simulations. Dans cette figure, comme dans toutes les figures analogues qui vont suivre, l'axe des ordonnées qui correspond à la valeur de P_{KS} est inversé et en échelle logarithmique : par conséquent, une faible valeur de P_{KS}, qui traduit une forte sensibilité, est matérialisée par une barre haute et inversement.

	Nombre de simulations utilisées					
Variables	**10 000 (1)**	**10 000 (2)**	**10 000 (3)**	**20 000**	**50 000**	**100 000**
ARGI	***	***	***	***	***	***
NORG	***	***	***	***	***	***
CALC	***	**	**	***	***	**
ALBEDO	*	*	*	*	*	*
q0	***	*	**	***	*	***
OBSTARAC	***	***	***	***	***	***
HCCF(1, 2)	***	***	***	***	***	***
HCCF(3)	***	**	**	***	***	***
HCCF(4)	*	*	*	*	*	*
HCCF(5)	*	*	*	*	*	*
HMINF(1, 2, 3, 5)	*	*	*	*	*	*
HMINF(4)	**	*	*	**	**	**
DA(1)	**	*	*	*	*	*
DA(2)	**	***	***	***	***	***
DA(3, 4)	*	*	*	*	*	*
DA(5)	*	*	*	*	**	*
NO3initf(1)	**	**	***	*	***	***
NO3initf(2, 3, 4, 5)	*	*	*	*	*	*
lvfront	*	*	*	*	*	*
durvieF	***	***	***	***	***	***
$vlai_{max}$	***	***	***	***	***	***

Tableau III.3-2 : Comparaison de l'effet du nombre de simulations initiales lors de l'application de la première itération de la méthode MCIP sur la sensibilité de la qualité des simulations aux différentes variables d'entrée testées. Pour 10 000 simulations initiales, trois séries ont été réalisées, notées (1), (2) et (3). *** indique une sensibilité maximale (P_{KS}<0.001) ; ** une sensibilité moyenne (0.001<P_{KS}<0.05) et * une faible sensibilité (P_{KS}>0.05). Année climatique utilisée : 1999-2000 ; information *a priori* : intervalles larges ; nombre d'itération : une. Les variables indiquées en gras sont celles qui ont été retenues.

D'après le Tableau III.3-2, le nombre de simulations initiales de la méthode MCIP joue assez peu sur les résultats de l'analyse de sensibilité. Certaines variables sont systématiquement déterminées sensibles, comme *ARGI*, *NORG*, *OBSTARAC*, *HCCF(1)*, *HCCF(2)*, *DA(2)*, *durvieF* et $vlai_{max}$. D'autres le sont de façon moins systématique : *CALC*, *q0*, *HCCF(3)* et *NO3initf(1)*. On peut, très schématiquement, classer ces variables selon

qu'elles influencent les simulations par leur rôle sur les bilans hydrique et azoté de la plante ou les deux ou encore sur la mise en place du système foliaire :

- bilan hydrique de la plante : *HCCF(h)*, *DA(2)*, déterminant la réserve utile ; *ARGI*, *q0* et *CALC*, jouant sur l'évaporation du sol ; *durvieF* et $vlai_{max}$ jouant sur le *LAI* et donc sur l'évapotranspiration ;

- bilan azoté de la plante : *NORG*, *NO3initf(1)*, *ARGI*, *CALC*, jouant sur la quantité d'azote disponible dans le sol ;

- les deux bilans : *OBSTARAC* qui limite l'espace exploré par les racines et *DA(2)* par l'influence sur les développement des racines ;

- mise en place du système foliaire : *durvieF* et *vlaimax*.

La mise en évidence de la sensibilité du modèle à ces variables est assez conforme à ce que l'on pouvait attendre d'après Ruget *et al.* (2002). L'apparition de certaines d'entre elles, comme *ARGI* et CALC, est néanmoins quelque peu surprenante. On aurait pu s'attendre à ce que pour une année humide comme 2000, les variables régulant l'évaporation du sol se révèlent moins importantes. En revanche, les humidités au point de flétrissement permanent, *HMINF(h)*, qui ne jouent qu'en condition de stress hydrique, ne ressortent pas de la présente analyse, conformément aux attentes. Il sera intéressant de confirmer cette analyse avec une autre année climatique, 1976. A ce stade, la liste des variables susceptibles d'être retenue pour la suite est donc : *ARGI*, *CALC*, *DA(2)*, *HCCF(1)*, *HCCF(2)*, *HCCF(3)*, *NO3initf(1)*, *NORG*, *OBSTARAC*, *q0*, *durvieF* et $vlai_{max}$ (indiquées en gras dans le Tableau III.3-2).

III.3.3.1.3 Influence du nombre d'itérations de la méthode sur l'analyse de sensibilité

Six itérations de la méthode MCIP ont été enchaînées. La Figure III.3-6 représente les résultats pour les trois premières itérations et la Figure III.3-7 sur les trois dernières. Au cours des diverses itérations de la méthode, au fur et à mesure que les simulations sont contraintes, de nouvelles variables d'entrée se révèlent sensibles au sens où elles permettent à leur tour d'améliorer la simulation de *LAI* et de *QN*. *DA(3)* apparaît ainsi dès la deuxième itération, tout comme *lvfront* et la *NO3initf(2)*. *HCCF(4)* apparaît à la quatrième itération, *DA(1)* à la cinquième. D'autres apparaissent de façon transitoire, comme *HPF(2)* et *HPF(4)*. Enfin, des variables jugées sensibles par les premières itérations disparaissent aux dernières itérations

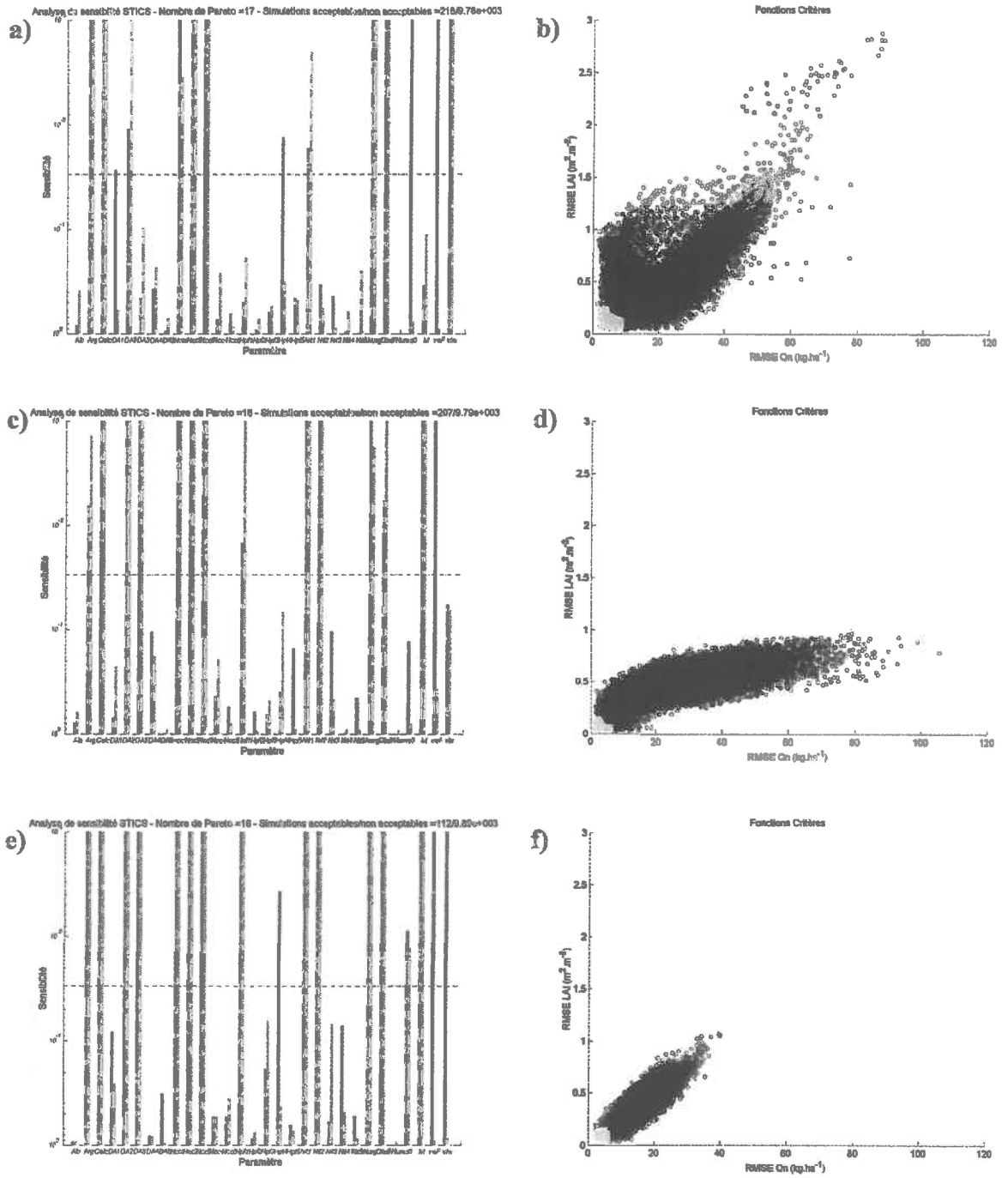


Figure III.3-6 : Analyse de sensibilité par la méthode MCIP sur les trois premières itérations (en lignes). Nombre de simulations utilisées : 10 000 ; information *a priori* : intervalles larges ; année climatique utilisée : 1999-2000. Les barres noires représentent les sensibilités multicritères, les grises, mono-critères. Les graphes b), d) et f) représentent les conséquences de la sélection et de la réduction progressive des intervalles de variation des variables d'entrée sur les valeurs des RMSE de *LAI* et de *QN*. En noir, l'ensemble des simulations réalisées à l'itération considérée, en gris, les simulations « acceptables ».

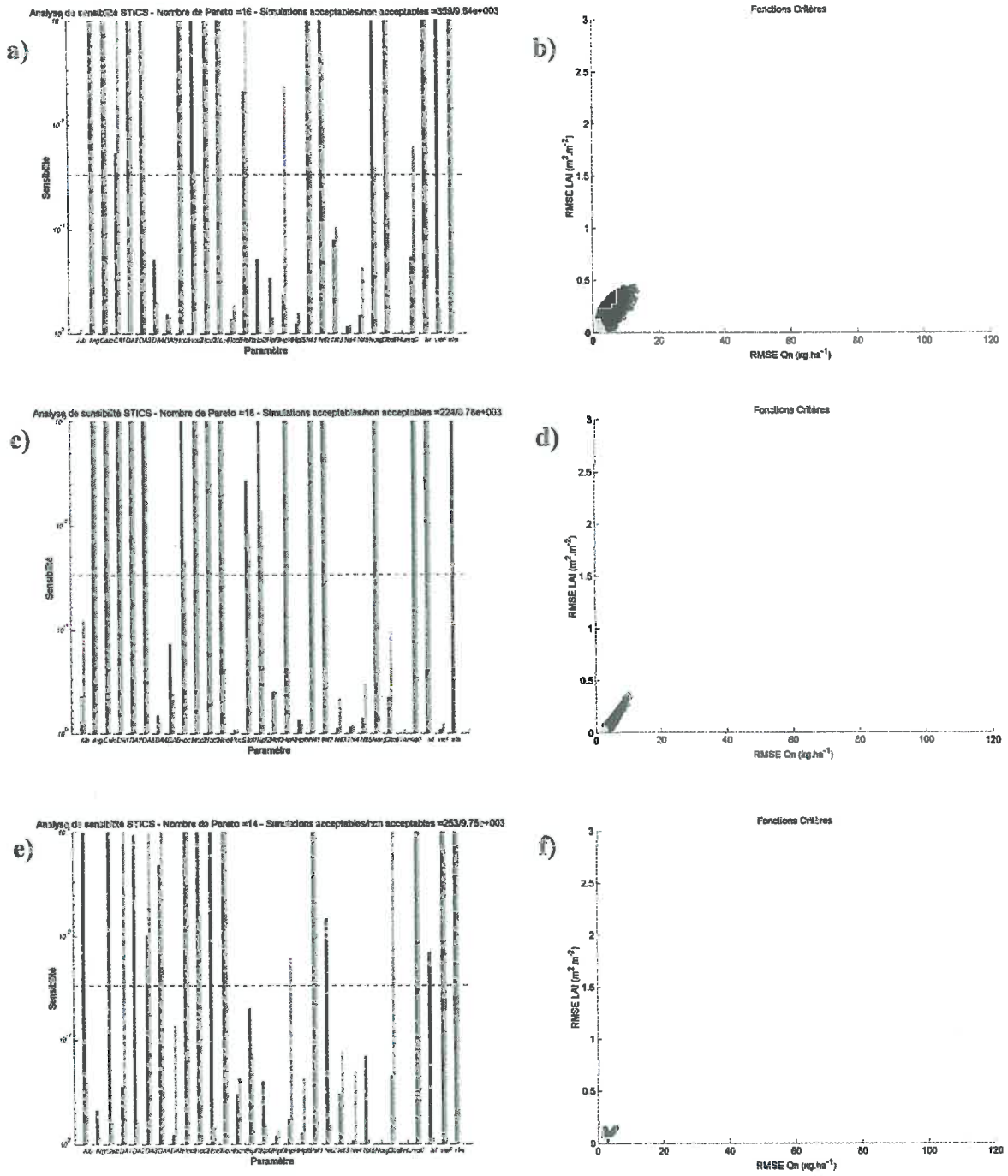


Figure III.3-7 : Analyse de sensibilité par la méthode MCIP sur les trois dernières itérations (en lignes). Nombre de simulations utilisées : 10 000 ; information *a priori* : intervalles larges ; année climatique utilisée : 1999-2000. Les barres noires représentent les sensibilités multicritères, les grises, mono-critères. Les graphes b), d) et f) représentent les conséquences de la sélection et de la réduction progressive des intervalles de variation des variables d'entrée sur les valeurs des RMSE de *LAI* et de *QN*. En noir, l'ensemble des simulations réalisées à l'itération considérée, en gris, les simulations « acceptables ».

après avoir joué leur rôle : *ARGI*, *NORG* et *OBSTARAC*. Mais dans la grande majorité des cas, on observe une certaine stabilité de la sensibilité des variables, au moins jusqu'à l'itération 3. Après celle-ci, les graphiques de la colonne de droite indiquent que les conséquences sur l'amélioration de *LAI* et de *QN* (traduite par la diminution de leur RMSE) sont de moindre portée. Les variables dont la sensibilité n'apparaît qu'aux itérations les plus tardives peuvent correspondre à des variables dont les effets sont masqués par d'autres. Ceci est lié aux compensations d'effet qui existent entre celles-ci.

Ainsi, la question se pose de savoir si aux douze variables retenues par la première itération, il ne serait pas nécessaire de rajouter au moins *DA(3)*, *NO3initf(2)* et *lvfront*. Ceci n'a pas été fait : l'effet des variables *DA(3)* et *NO3initf(2)* peut être compensé par *DA(2)* et par *NO3initf(1)* et celui de *lvfront*, par le rôle de la densité apparente, traduit par *DA(2)*, sur la simulation de l'enracinement.

III.3.3.1.4 Influence de la qualité de l'information *a priori* sur l'analyse de sensibilité

L'effet de la nature de l'information *a priori* a été envisagé en réalisant des analyses utilisant deux types d'intervalles entre lesquels sont générés les $N_M = 10\,000$ vecteurs de variables d'entrée : des « intervalles étroits », encadrant de ±10% les valeurs utilisées dans la simulation de référence et des « intervalles larges », correspondant à une information *a priori* sur la parcelle. Les résultats pour les intervalles étroits sont indiqués par la Figure III.3-8 et sont à comparer à ceux obtenus pour les intervalles larges déjà présentés dans la Figure III.3-5.

La qualité de l'information *a priori* affecte la notion de sensibilité au sens de la méthode MCIP : seules *DA(2)*, *DA(3)*, *durvieF* et $vlai_{max}$ ressortent de l'analyse de sensibilité, *NORG* étant moyennement sensible. Lorsque la connaissance de la valeur d'une variable est bonne, la qualité des simulations n'est pas sensible à cette dernière au sens où le fait de changer sa valeur entre des intervalles restreints ne permet pas d'améliorer les simulations par rapport aux observations. Tout se passe finalement comme si un certain nombre d'itérations avaient déjà été effectuées. Ce point montre la nécessité de faire une analyse de sensibilité de ce type chaque fois qu'on change d'information *a priori* : ceci revient à dire la trivialité qui veut que si la valeur d'une variable est bien connue, il ne sert à rien de chercher à l'estimer. La méthode MCIP permet de faire ce tri de façon automatique. Il s'agit cependant de garder à l'esprit que le type d'information *a priori* « borne étroite » est purement conceptuel.

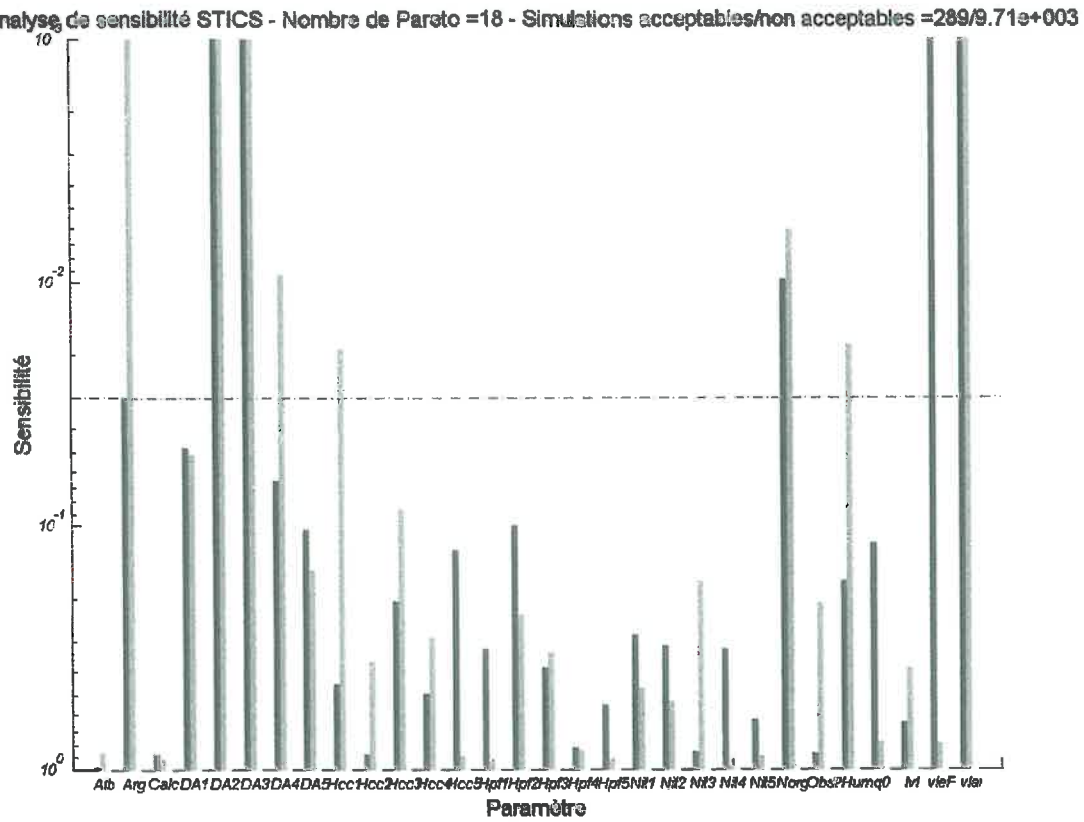


Figure III.3-8 : Analyse de sensibilité par la méthode MCIP en utilisant les intervalles étroits (proches de la valeur de référence) pour chaque variable d'entrée. Nombre de simulations utilisées : 10 000 ; année climatique utilisée : 1999-2000 ; nombre d'itération : une. Les noms des variables d'entrée présentés en abscisse sont respectivement : *ALBEDO*, *ARGI*, *CALC*, *DA(1)*, *DA(2)*, *DA(3)*, *DA(4)*, *DA(5)*, *HCCF(1)*, *HCCF(2)*, *HCCF(3)*, *HCCF(4)*, *HCCF(5)*, *HMINF(1)*, *HMINF(2)*, *HMINF(3)*, *HMINF(4)*, *HMINF(5)*, *NO3initf(1)*, *NO3initf(2)*, *NO3initf(3)*, *NO3initf(4)*, *NO3initf(5)*, *NORG*, *OBSTARAC*, *PROFHUM (neutralisée)*, *q0*, *lvfront*, *durvieF* et $vlai_{max}$. Les barres noires représentent les sensibilités multicritères, les grises, mono-critères.

III.3.3.1.5 Influence du climat sur l'analyse de sensibilité

Deux années climatiques ont été utilisées afin de vérifier l'effet éventuel du climat sur la sensibilité de la qualité des simulations aux variables d'entrées : 1975-1976, caractérisée par une importante période de sécheresse et 1999-2000, beaucoup plus humide, qui est une des années du jeu de données de Chambry. Les résultats obtenus pour 1975-1976 sont présentés dans la Figure III.3-9. Ceux de 1999-2000 ont déjà été présentés dans la Figure III.3-5.

Quelques changements de sensibilité ont lieu entre les deux années climatiques utilisées. La plupart des variables qui étaient jugées sensibles le restent (*ARGI*, *CALC*, *DA(2)*, *HCCF(1)*, *HCCF(2)*, *HCCF(3)*, *NO3initf(1)* et *OBSTARAC*). Mais les humidités au point de

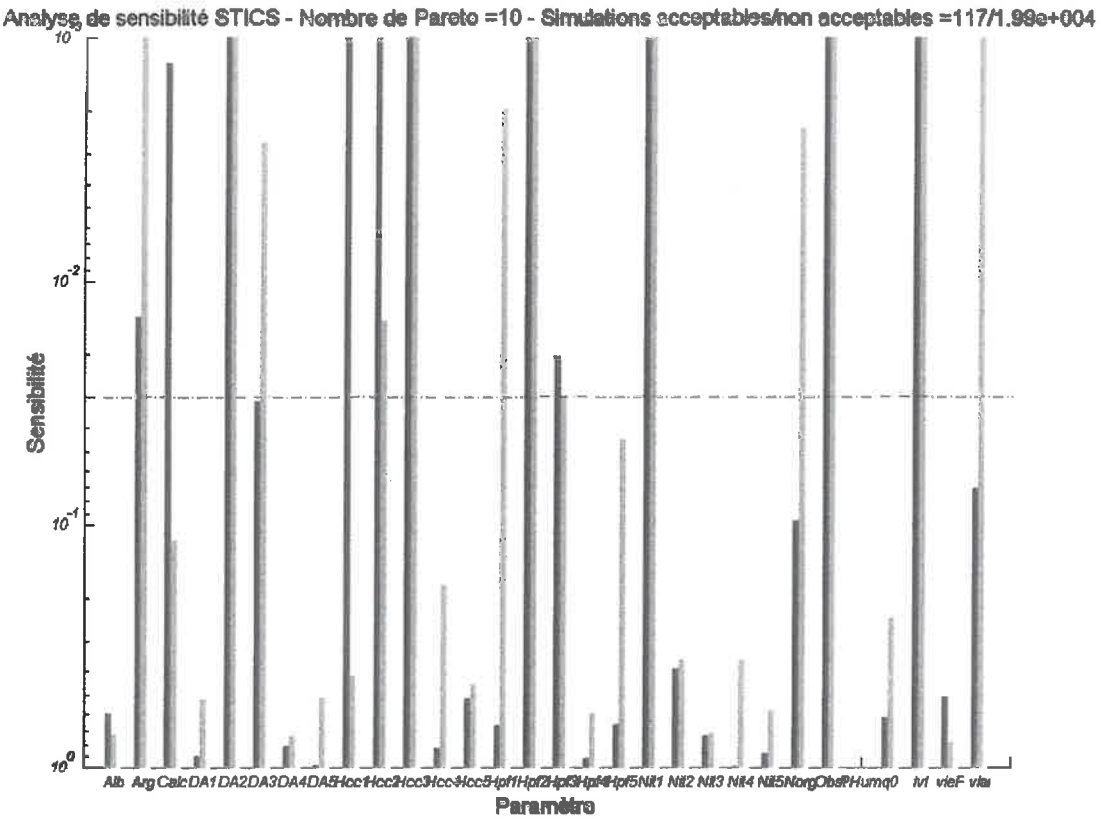


Figure III.3-9 : Analyse de sensibilité par la méthode MCIP en utilisant l'année climatique 1975-1976. Nombre de simulations utilisées : 10 000 ; information *a priori* : intervalles larges ; nombre d'itération : une. Les noms des variables d'entrée présentés en abscisse sont respectivement : *ALBEDO*, *ARGI*, *CALC*, *DA(1)*, *DA(2)*, *DA(3)*, *DA(4)*, *DA(5)*, *HCCF(1)*, *HCCF(2)*, *HCCF(3)*, *HCCF(4)*, *HCCF(5)*, *HMINF(1)*, *HMINF(2)*, *HMINF(3)*, *HMINF(4)*, *HMINF(5)*, *NO3initf(1)*, *NO3initf(2)*, *NO3initf(3)*, *NO3initf(4)*, *NO3initf(5)*, *NORG*, *OBSTARAC*, *PROFHUM (neutralisée)*, *q0*, *lvfront*, *durvieF* et $vlai_{max}$. Les barres noires représentent les sensibilités multicritères, les grises, mono-critères.

flétrissement permanent des horizons 2 et 3 apparaissent, ce qui est le reflet des conditions sèches de cette année-là. En corollaire, la variable *NORG* n'est plus sensible, ce qui montre que dans les conditions de cette année, la disponibilité de l'azote était un facteur moins limitant que la disponibilité en eau. Par contre, fait étonnant, *q0*, qui intervient pourtant dans le calcul de l'évaporation du sol, n'apparaît plus. On peut supposer que son effet est soit masqué par celui de *HMINF(2)* et *HMINF(3)*, soit qu'il intervient plus tôt, en début de cycle, alors que la sécheresse ne s'est pas encore déclarée. A ce moment-là, comme c'est surtout cette période qui détermine la sensibilité des variables, son effet devient mineur. Ceci illustre le fait qu'il existe une relation entre la sensibilité à une variable et l'existence d'un facteur limitant associé à cette variable au moment où la sensibilité est étudiée. Pour cette raison, les différentes dates d'observation ne jouent pas le même statut.

Un autre fait intéressant est l'inversion des rôles des paramètres plantes : *lvfront* qui n'avait pas été jugé sensible jusqu'à présent le devient, et *durvieF* et $vlai_{max}$ ne sont pas ici sélectionnés. Ceci montre, en condition de sécheresse, la prépondérance du développement racinaire sur le fonctionnement de la plante.

Ces constatations suggèrent au moins deux conclusions partielles :

- la nécessité de réaliser une analyse de sensibilité par année climatique ;

- le fait, sur lequel nous reviendrons, que pour estimer correctement toutes les variables caractéristiques du sol traduisant son fonctionnement vis à vis des bilans hydriques et azotés, une seule année climatique ne suffit pas, car elle ne permet pas d'exprimer tous les facteurs limitants. En corollaire, les valeurs estimées une année donnée seront valables pour cette année, pas forcément pour d'autres.

III.3.3.1.6 Conclusion : séiection des variables d'entrée pour le reste de l'étude

Les variables d'entrée retenues pour effectuer la comparaison des trois méthodes sont celles qui ont été choisies dans la partie III.3.3.1.2 p. 179 : *ARGI*, *CALC*, *DA(2)*, *HCCF(1)*, *HCCF(2)*, *HCCF(3)*, *NO3initf(1)*, *NORG*, *OBSTARAC*, *q0*, *durvieF* et $vlai_{max}$. On a vu néanmoins que des variables comme *DA(3)*, *NO3initf(2)* et *lvfront* auraient pu également être prises en compte.

En ce qui concerne le nombre de simulations à réaliser avec la méthode MCIP, nous en resterons, au vu des résultats, à 10 000.

Retenir douze variables d'entrée à estimer peut paraître un nombre trop important. Mais ceci est rendu possible par le fait qu'on utilise une information *a priori* sur chacun d'entre eux. En outre, ceci permet d'explorer une plus large gamme de combinaison de ces variables d'entrée, et donc une plus grande variabilité des variables d'état pour lesquelles on a des observations. Avec moins de variables d'entrée, pour atteindre cette souplesse, il faudrait donner des gammes de variation aux variables d'entrée qui pourraient ne plus rien à voir avec leurs gammes de variation plausibles.

Les conclusions concernant la mise en pratique de cette méthode d'analyse de sensibilité sont qu'elle doit être appliquée dès lors que le contexte d'application change, soit que (i) l'information *a priori* change de précision, ce qui remet en cause le poids des différentes variables d'entrée pour contraindre les simulations ; (ii) soit que la situation

agronomique, pédologique ou climatique change, modifiant la hiérarchie des facteurs limitants et donc celle des variables d'entrée influentes. L'étude de l'effet du jeu de variables définissant la situation de référence n'a pas été envisagé dans cette étude. Il aurait fallu pour ce faire répéter semblable analyse pour différents jeux de valeurs. D'un point de vue pratique, il apparaît qu'il faudrait sélectionner les variables sensibles pour tous les points où l'on veut réaliser une assimilation, tout comme on doit les sélectionner pour tous les climats. Mais, pour la comparaison des trois méthodes d'assimilation qui va suivre, il était plus simple et plus cohérent d'estimer toujours les mêmes variables.

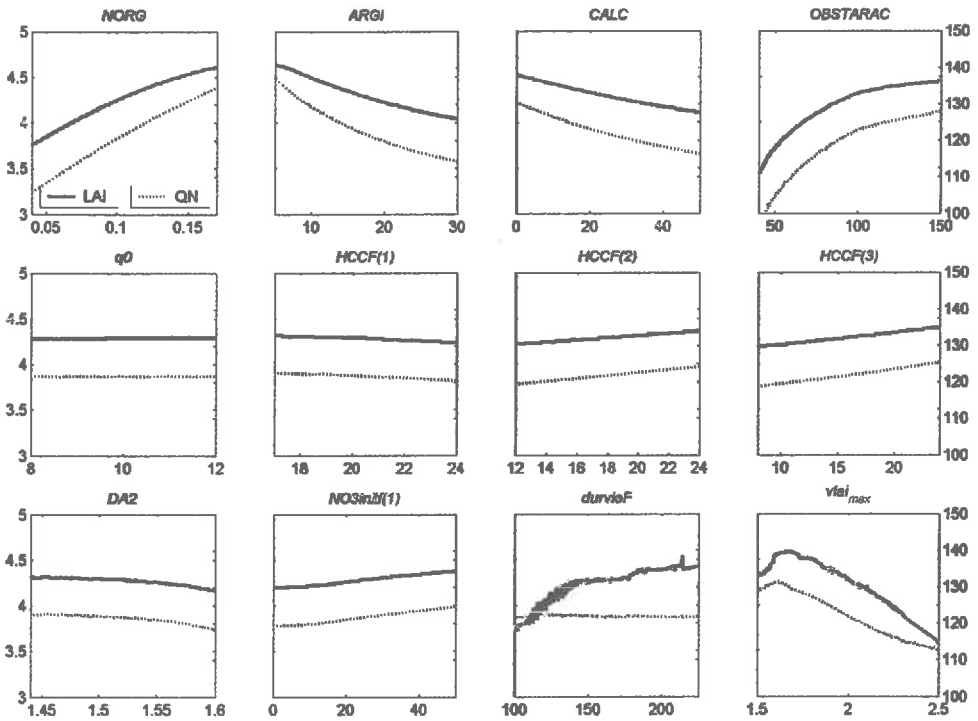


Figure III.3-10 : Evolution des variables *LAI* et de *QN* simulées au 3 mai pour une variation d'une seule des 12 variables d'entrée sélectionnées. Les intervalles de variation correspondent aux « intervalles larges ». Les valeurs des variables de référence correspondent au centre de l'intervalle.

La Figure III.3-10 représente la variation des variables *LAI* et *QN* simulées au 3 mai en fonction de la valeur de chacune des 12 variables d'entrée considérées isolément, les autres étant fixées au centre des intervalles larges. La Figure III.3-11 fait de même avec le rendement, *G* et la teneur en protéines, *P*. On constate que généralement les variables qui jouent un rôle fort sur *LAI* et *QN* ont également une influence marquée sur le rendement et la teneur en protéines et que pour les valeurs de référence choisies, alors que certaines variables comme *q0* et *HCCF(1)* ont assez peu d'impact sur *LAI* et *QN* (du moins au 3 mai). *DA(2)* a très peu d'effet sur les variables à la récolte. Mais cette approche ne tient pas compte des

interactions entre les différentes variables d'entrée : elle a ici plus une vertu illustrative de la façon dont une variable d'entrée donnée, les autres étant fixées, peut influencer les différentes variables simulées qui nous intéressent. On remarque également un comportement pour le moins erratique de G et P en fonction de certaines variables d'entrée. Ces instabilités numériques, liées au code de STICS, nécessiteraient par elles-mêmes une révision complète de sa programmation et peuvent être à l'origine de nombreux déboires essuyés au cours des essais d'assimilation de données issues de la télédétection par estimation des entrées spatialement variables du modèle (*cf. infra*).

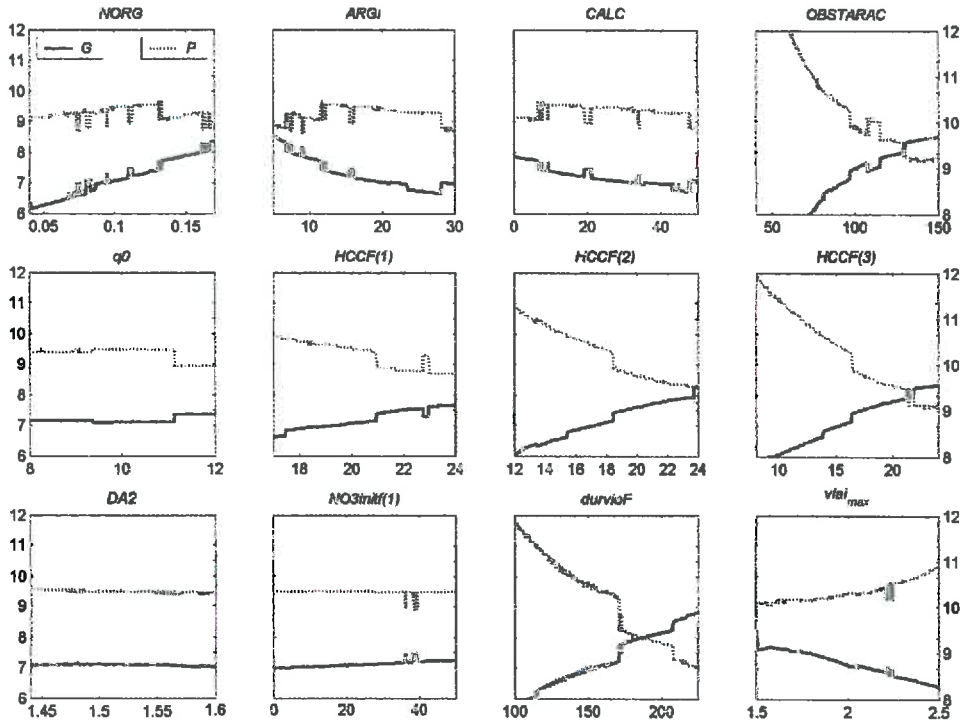


Figure III.3-11 : Evolution du rendement, G et de la teneur en protéines, P, pour une variation d'une seule des 12 variables d'entrée sélectionnées. Les intervalles de variation correspondent aux « intervalles larges ». Les valeurs des variables de référence correspondent au centre de l'intervalle.

III.3.3.2 Protocole de comparaison des trois méthodes

Une fois déterminées les variables d'entrée du modèle sur lesquelles doit porter l'assimilation, une comparaison des méthodes SIMPLEXE, GLUE et MCIP a été réalisée afin de sélectionner celle qui sera utilisée sur le jeu de données de Chambry. Le protocole de comparaison utilisé est synthétisé par la Figure III.3-12. Après la définition des objectifs, cette partie comprend trois phases : (i) la création du jeu de données nécessaire ; (ii) la description des paramètres d'utilisation des trois méthodes ; (iii) les critères de comparaison choisis.

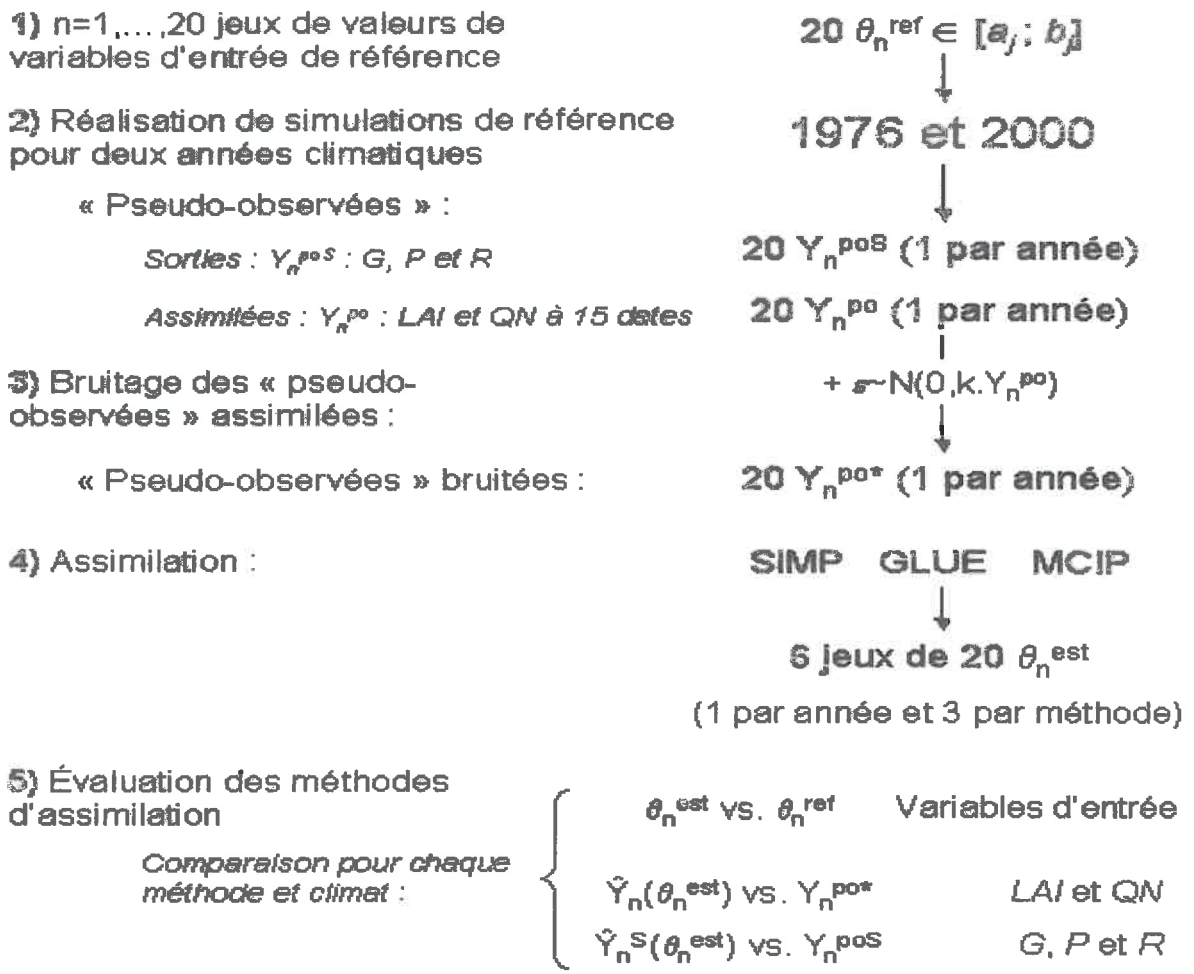


Figure III.3-12 : Représentation schématique du protocole de comparaison des méthodes d'assimilation SIMPLEXE, GLUE et MCIP.

III.3.3.2.1 Définition des objectifs

Le choix de travailler sur des jeux de données simulées a été guidé par la volonté de comparer les trois méthodes d'assimilation en faisant abstraction des difficultés pouvant provenir du modèle, typiquement, ses erreurs de structure. On se place ainsi dans une situation où le modèle décrit parfaitement la réalité et où les erreurs ne peuvent provenir que des incertitudes sur les variables d'entrée ou les observations. Vingt situations de référence correspondant chacune à une zone homogène d'une parcelle et caractérisées par des valeurs caractéristiques du sol ont été simulées.

Deux types de variables d'entrée sont à considérer : (i) celles qui sont estimées par assimilation : pour chacune d'elle, on a supposé connue une information *a priori* (IAP) ; (ii) celles qui ne sont pas estimées et qui auront la même valeur pour créer le jeu de données et lors de l'assimilation. Afin de reproduire l'effet de la méconnaissance de ces dernières sur l'assimilation, elles auraient pu être bruitées, mais la façon de réaliser ce bruitage aurait pu conditionner les résultats de la comparaison.

Afin d'évaluer l'effet du climat sur l'assimilation, deux années ont été employées, les mêmes que celles utilisées pour l'analyse de sensibilité : 1975-1976 et 1999-2000.

Les trois méthodes d'assimilation seront comparées sur des critères correspondant à des objectifs appliqués : (i) leur capacité à décrire les variables d'entrée sol qui ont servi à faire les simulations de référence ; (ii) leur capacité à trouver des variables d'entrée qui permettent de bien simuler les variables d'intérêt agronomique : le rendement, la teneur en protéines et le reliquat d'azote. Ces capacités seront évaluées par rapport à la description permise par l'IAP.

III.3.3.2.2 Création des jeux de données de référence

Etape 1 de la Figure III.3-12 – 20 simulations de référence ont été réalisées à partir de 20 jeux comprenant les variables d'entrées citées dans le Tableau III.3-3 pour les deux années climatiques. Afin d'avoir des combinaisons réalistes des valeurs de ces variables, ces valeurs de référence (θ_n^{ref} avec n=1,…,20) ont été extraites de la base de données cartographique de chacune des deux parcelles expérimentales, dans des zones contrastées. Les bornes définissant l'IAP ont également été fixées d'après cette base de données : elles correspondent aux valeurs minimales et maximales observées pour chaque variable. Les bornes inférieure et supérieures des paramètres plante ont été fixées de façon à encadrer les valeurs prises dans la version standard : 2 pour $vlai_{max}$ et 165 pour *durvieF*.

Etape 2 de la Figure III.3-12 – Une correspondance a ainsi été établie entre 20 vecteurs de 12 variables d'entrée θ_n^{ref} et 20 vecteurs de 30 variables d'état « pseudo-observées », composé des quinze observations de *LAI* et des quinze de *QN*. Les dates d'observation sont, comme pour l'analyse de sensibilité : 16/02/00, 01/03/00, 15/03/00, 29/03/00, 05/04/00, 12/04/00, 19/04/00, 26/04/00, 03/05/00, 10/05/00, 17/05/00, 24/05/00, 07/06/00, 21/06/00, 05/07/00. On associe également aux 20 vecteurs θ_n^{ref} 20 vecteurs composés des variables de

sortie d'intérêt final, qui serviront à juger de l'intérêt de l'assimilation vis-à-vis de la restitution de ces variables.

Variable	STICS	Borne inférieure	Borne supérieure
Teneur en azote organique (H1)	*NORG* (%p)	0.04	0.17
Teneur en argile après décarbonatation (H1)	*ARGI* (%p)	5	30
Teneur en calcaire (H1)	*CALC* (%p)	0	50
Obstacle physique ou chimique	*OBSTARAC* (cm)	40	150
Paramètre d'évaporation (H1)	q0 (mm)	8	12
Humidité à la capacité au champ (H1)	*HCCF(1)* (%p)	17	24
Humidité à la capacité au champ (H2)	*HCCF(2)* (%p)	12	24
Humidité à la capacité au champ (H3)	*HCCF(3)* (%p)	8	24
Densité apparente (H2)	*DAF(2)* (g cm-3)	1.44	1.6
Azote au semis (H1)	*NO3initf(1)* (kg ha-1)	0	50
Durée de vie des feuilles (variable plante)	*durvieF*	100	225
Etablissement du *LAI* (variable plante)	$vlai_{max}$	1.5	2.5

Tableau III.3-3 : Valeurs minimales et maximales utilisées comme information *a priori* pour chaque variable d'entrée du modèle utilisée dans l'assimilation (%p désigne un pourcentage pondéral. La lettre H suivie d'un chiffre désigne le numéro de l'horizon concerné).

Etape 3 de la Figure III.3-12 – Afin de simuler les erreurs sur les observations, nous avons bruité les « pseudo-observations ». Pour ce faire, un bruit additif a été ajouté aux variables pseudo-observées, notées Y_n^{po}* :

$\forall n \in [1 ; 20]$, $Y_n^{po*}=Y_n^{po}+\varepsilon_n$ Équation III.3-13

où Y désigne *LAI* ou *QN*, ε_n est un vecteur de variables aléatoires tirées suivant une loi normale :

$\varepsilon_n \sim N(0 ; \sigma^2)$ avec $\sigma = k \bullet Y_n^{po}$ Équation III.3-14

où k est un vecteur déterminé expérimentalement et contenant les valeurs k_{LAI}=0.09 pour *LAI* et k_{QN}=0.08 pour *QN*. Ces valeurs correspondent aux coefficients de variation de mesures de *LAI* et de *QN* sur des jeux de données observées.

III.3.3.2.3 Paramètres d'utilisation des trois méthodes

Etape 4 de la Figure III.3-13 – Dans les critères à minimiser dans les trois méthodes, le rôle des valeurs observées est donc tenu par les valeurs pseudo-observées bruitées, $Y_n^{po}*$. Les valeurs des erreurs d'observations dans les différents critères à minimiser sont $\sigma_{LAI}^{vo} = k_{LAI} \bullet LAI_n^{po}$ et $\sigma_{QN}^{vo} = k_{QN} \bullet QN_n^{po}$ et le nombre de dates d'observation est D=15.

Pour la méthode SIMPLEXE, chaque assimilation a été réalisée avec 25 vecteurs initiaux tirés de façon uniforme dans l'IAP. Pour la méthode GLUE, N_G = 200 000 simulations de départ ont été réalisées par année climatique en tirant aléatoirement selon une loi uniforme les valeurs des variables d'entrée dans les intervalles définis par l'IAP. Ces simulations servent pour les 20 jeux de données car on considère que l'IAP est commune à ces 20 jeux. Pour la méthode MCIP, les deux fonctions objectifs à optimiser simultanément étaient des RMSE pondérées calculées sur les valeurs de *LAI* et de *QN*. Six itérations ont été utilisées pour chacun des 20 jeux de données, avec à chaque fois 10 000 simulations effectuées. Les assimilations effectuées ont permis de déterminer des vecteurs (SIMPLEXE) ou des distributions *a posteriori* (GLUE et MCIP) des variables d'entrée estimées, notées θ_n^{est}. On a autant de valeurs de θ_n^{est} que d'années et de méthodes utilisées, soit six en tout.

Etape 5 de la Figure III.3-12) – Une fois θ_n^{est} déterminés, des simulations ont été réalisées à partir des distributions obtenues. Pour la méthode SIMPLEXE, une seule simulation a été effectuée pour chacun des 20 jeux de θ_n^{est} ; pour les méthodes GLUE et MCIP, 1000 simulations ont été effectuées pour chacun des 20 jeux en tirant aléatoirement les valeurs de variables d'entrée dans la distribution de θ_n^{est} obtenue. Pour caractériser le niveau de connaissance sans assimilation, 1000 simulations ont été effectuées en tirant les valeurs des variables d'entrée selon une loi uniforme définie par l'IAP. La moyenne de ces 1000 simulations servira de référence.

III.3.3.2.4 Critères de comparaison des trois méthodes

Les performances des trois méthodes ont été comparées selon les critères suivants (étape 5 de la Figure III.3-12) :

- leur capacité à retrouver la valeur des variables d'entrée ayant servi à réaliser les simulations de référence, θ_n^{ref}, jugée par la RMSE entre θ_n^{ref} et la moyenne de la distribution de θ_n^{est} ;

- leur capacité à restituer des valeurs de variables d'entrée qui permettent de bien simuler les variables servant à l'assimilation, *LAI* et *QN*. Elle sera évaluée à travers la RMSE calculée sur *LAI* et *QN*. Comme les mêmes dates d'observation seront utilisées, il ne s'agira pas à proprement parler d'une validation de la méthode mais plutôt d'une vérification de son efficacité ;

- leur capacité à déterminer des valeurs de variables d'entrée qui permettent de bien restituer les variables d'intérêt, rendement (*G*), teneur en protéines (*P*), reliquat d'azote à la récolte (*R*). Ce critère fera office de validation puisque les variables considérées n'auront pas été utilisées pour réaliser l'assimilation. Cette capacité sera évaluée à travers la RMSE calculée sur ces variables de sortie ;

- le nombre de simulations à réaliser.

Les différentes RMSE ont été calculées entre les valeurs des jeux de référence (θ_n^{ref} et Y_n^{po}* et Y_n^{poS}) et les moyennes des distributions déterminées par assimilation (θ_n^{est}, $\hat{Y}_n(\theta_n^{est})$ et $\hat{Y}_n^S(\theta_n^{est})$ et distribution de l'IAP). Les RRMSE, RMSEs et RMSEu sont calculées entre les mêmes vecteurs.

III.3.3.3 Résultats de la comparaison

III.3.3.3.1 Capacité à retrouver (θ_n^{ref})

Pour chacun des 20 vecteurs de variables d'entrée de référence, θ_n^{ref}, on a trouvé six ensembles de 20 vecteurs de variables d'entrée estimés par assimilation, θ_n^{est}, trois méthodes différentes et deux années climatiques ayant été considérées. Pour chacun de ces six ensembles, on calcule des valeurs de RRMSE pour chaque variable d'entrée entre la valeur de référence et la valeur retrouvée (Tableau III.3-4). Afin de juger de l'amélioration des simulations permise par les différentes méthodes, la RRMSE des variables d'entrée renseignées d'après l'information *a priori* est également donnée (IAP).

On constate tout d'abord que les RRMSE de l'IAP ont des valeurs très diverses : ceci illustre le fait que, en conditions réelles, la qualité de l'IAP n'est pas la même pour toutes les variables d'entrée. Pour $vlai_{max}$, vu que la valeur de référence était de 2 pour toutes les simulations de référence et que l'IAP était [1.5 ; 2.5], la RRMSE de l'IAP est nulle (*cf.* Tableau III.3-3). Pour le calcaire, l'IAP est [0 ; 50] mais les valeurs de référence se situaient

plutôt dans la première partie de cette gamme, et c'est ce qui explique la forte valeur de RRMSE de l'IAP.

		1976	2000	1976	2000	1976	2000
Variable	IAP	SIMPLEXE	SIMPLEXE	GLUE	GLUE	MCIP	MCIP
NORG	20.2	44	41	26.9	23.5	31.2	24.8
ARGI	37.4	68	56	34.8	36	44.6	30.8
CALC	164	198	499	140	154	194	195.1
OBSTARAC	33.4	42	52	17.3	28.3	22.7	37.3
q0	12.7	20	17	14	11.2	18.7	15.3
HCCF(1)	8.7	25	15	7.6	7.4	13	9.86
HCCF(2)	13.5	37	22	12.9	12.7	14.8	18.5
HCCF(3)	21.6	32	33	22.1	24.6	33	35.6
DA2	3.2	6	5	3.18	3.09	4.54	3.57
$NO3_{init}1$	70.1	102	97	68.5	70.6	89.9	94.4
durvieF	11.2	22	9	6.33	2.29	6.69	1.56
$vlai_{max}$	0	17	12	1.18	1.65	1.08	1.85

Tableau III.3-4 : Valeurs des RRMSE de chaque variable d'entrée retrouvée par les différentes méthodes d'assimilation (θ_n^{est}) par rapport aux variables d'entrée ayant servi à faire les simulations de référence (θ_n^{ref}). La RRMSE obtenue à partir de l'information *a priori* (IAP) est donnée à titre de référence. 1976 : valeurs retrouvées en utilisant l'année climatique 1975-1976 ; 2000 : en utilisant l'année climatique 1999-2000.

La Figure III.3-13 présente ces valeurs de façon normalisée : la valeur de RRMSE de l'IAP est prise comme référence et est ramenée à 1, sauf pour la variable $vlai_{max}$ pour laquelle la valeur obtenue par la méthode SIMPLEXE de 1976 a été choisie comme référence.

Cette représentation permet de voir que les valeurs de variables d'entrée sont de manière générale très mal retrouvées. La méthode SIMPLEXE conduit toujours à des estimations moins bonnes que celles obtenues avec l'IAP. Les méthodes GLUE et MCIP donnent de ce point de vue de meilleurs résultats, la méthode GLUE conduisant dans la plupart des cas à la meilleure restitution des variables. Seules les variables *durvieF* et $vlai_{max}$ sont assez bien retrouvées avec des valeurs de RRMSE très faibles (pour les méthodes GLUE et MCIP). Ceci confirme la sensibilité des simulations du modèle à ces variables.

En ce qui concerne l'effet de l'année climatique sur la qualité de l'estimation de θ_n^{est}, on peut remarquer deux différences notables : *OBSTARAC* est bien mieux retrouvé avec la série de 1976 et *durvieF* avec celle de 2000. On aurait pu s'attendre à ce que le climat joue

d'une façon plus déterminante sur les performances d'estimation des valeurs des variables d'entrée. Cette influence semble donc limitée à quelques variables et montre que *OBSTARAC* est déterminant pour ce qui est de la simulation des stress hydriques (année 1976), tandis que

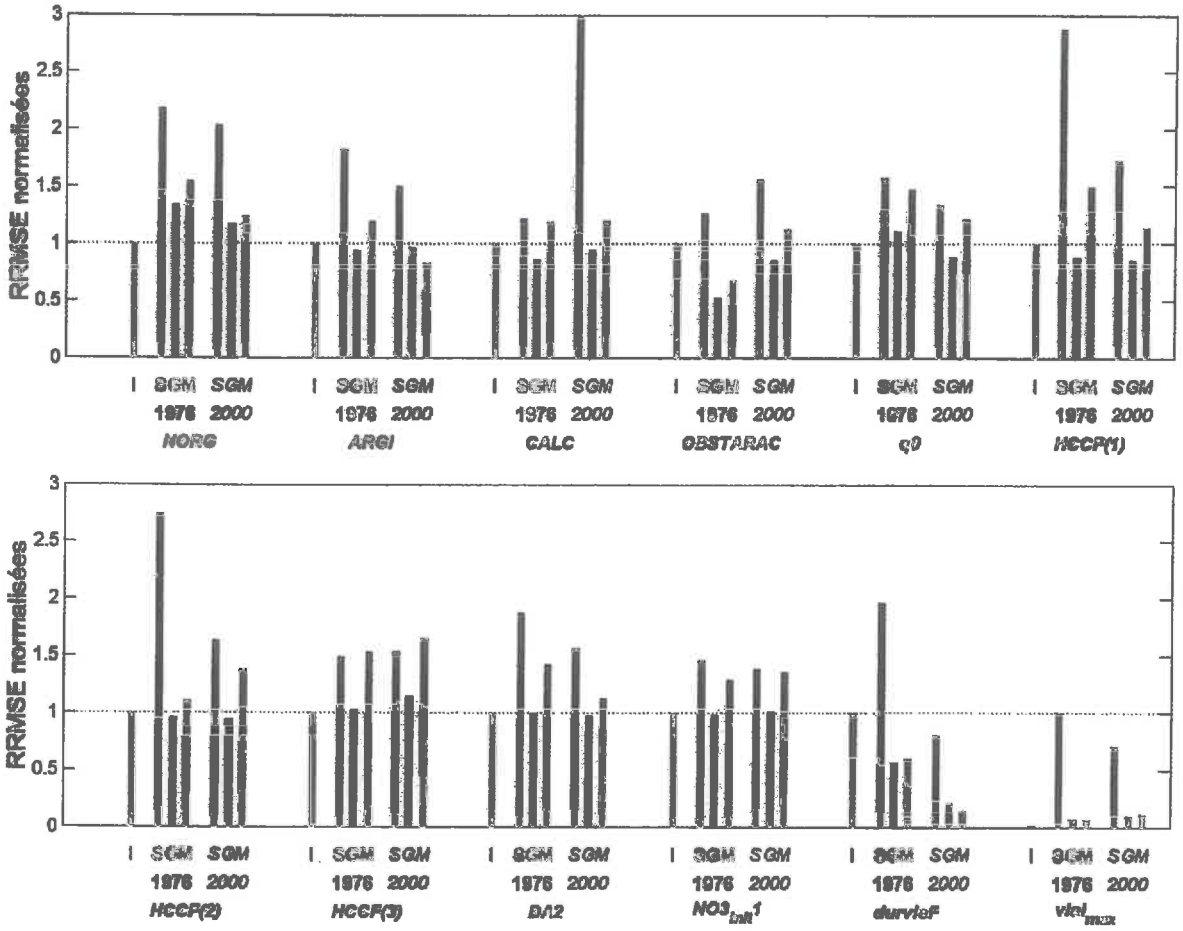


Figure III.3-13 : RRMSE entre θ_n^{est} et θ_n^{ref} normalisées entre 0 et 1 pour les années climatiques 1975-1976 (1976 sur la figure) et 1999-2000 (2000). I : valeur normalisée de la RRMSE des variables d'entrée connues d'après l'information *a priori* ; S : valeur normalisée de la RRMSE des variables d'entrée estimées par la méthode SIMPLEXE ; G : par la méthode GLUE ; M : par la méthode MCIP. Les valeurs de RRMSE ont été normalisées à 1 par rapport à l'information *a priori*, sauf dans le cas de $vlai_{max}$ pour laquelle celle-ci est nulle. C'est la valeur obtenue par la méthode SIMPLEXE pour 1976 qui a dans ce cas servi de référence. Pour les valeurs non normalisées, voir Tableau III.3-4.

durvieF est plus important dans des conditions moins stressantes. Quant à $vlai_{max}$, qui joue très directement sur la dynamique d'installation du *LAI*, il semble que les simulations dont la valeur s'éloigne trop de la valeur de référence sont systématiquement mauvaises, quel que soit le type d'année. Ceci confirme son importance et montre que si une mauvaise valeur lui est affectée, elle ne peut pas être compensée par les autres variables.

Ces résultats montrent qu'aucune des trois méthodes ne permet d'estimer correctement les valeurs des variables d'entrée du modèle de culture par assimilation. Ceci peut être directement attribué aux compensations qui peuvent exister entre les différentes variables d'entrée. En effet, le modèle n'étant pas bijectif, on peut imaginer que plusieurs jeux de valeurs de variables puissent donner des valeurs de *LAI* et *QN* proches. Cette hypothèse va être examinée en évaluant la qualité des simulations obtenues à partir de θ_n^{est} par rapport aux simulations de référence. Les instabilités numériques observées dans le modèle (Figure III.3-10 et Figure III.3-11) contribuent à créer des minima locaux artificiels et sont de ce fait très certainement, pour une part, à l'origine des ces mauvais résultats.

III.3.3.3.2 Capacité à restituer *LAI* et *QN*

La référence, comme pour les variables d'entrée, sera constituée des simulations réalisées à partir de l'IAP.

Les valeurs de RRMSE, RRMSEs et RRMSEu de *LAI* et *QN*, calculées sur l'ensemble des 15 dates et des 20 points, sont données dans la Figure III.3-14. On constate que toutes les assimilations améliorent la qualité de simulation de *LAI* et *QN* par rapport à l'IAP, mis à part la méthode SIMPLEXE pour l'année climatique 1976. La part systématique de l'erreur (estimée par RMSEs) est toujours faible (moins de 6%) : les valeurs retrouvées n'introduisent donc pas de biais dans les simulations mais plutôt une dispersion (évaluée par RMSEu) qui explique la majeure partie des erreurs commises.

Le fait que la méthode SIMPLEXE ne permet pas de simuler des valeurs correctes de *LAI* et *QN* ni de bien estimer les valeurs des variables d'entrée montre qu'il faudrait améliorer le protocole de sa mise en œuvre afin de mieux explorer l'espace des solutions. La qualité de l'estimation de *LAI* et *QN* par les méthodes GLUE et MCIP est bonne (aux alentours de 15% d'erreur), mais ce qu'il est plus important de vérifier pour valider ces méthodes est la qualité des simulations des variables d'intérêt, le rendement *G*, la teneur en protéines *P* et le reliquat d'azote à la récolte *R*, qui n'ont pas servi à l'assimilation.

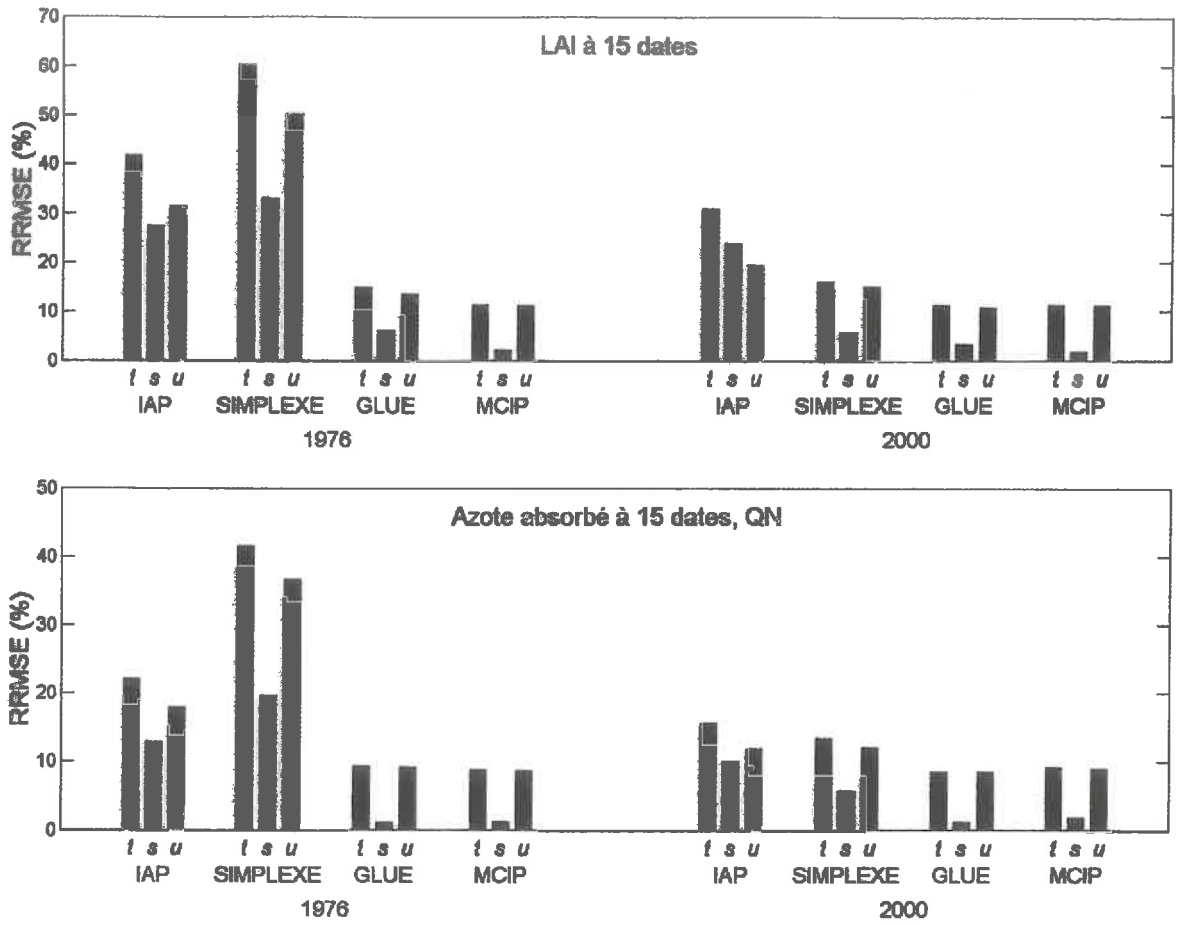


Figure III.3-14 : RRMSE entre les valeurs de *LAI* et *QN* « pseudo-observées » bruitées et les valeurs simulées à partir de l'IAP ou des variables d'entrées estimées (calculées sur les 20 points) pour les années climatiques 1975-1976 (1976 sur la figure) et 1999-2000 (2000). Les valeurs de RRMSE (*t*), RRMSEs (*s*) et RRMSEu (*u*) sont données pour les simulations réalisées d'après l'information *a priori* et d'après les valeurs de variables d'entrée obtenues par assimilation par les méthodes SIMPLEXE, GLUE et MCIP.

III.3.3.3.3 Capacité à restituer *G*, *P* et *R*

La Figure III.3-15 représente les RRMSE, RRMSEs et RRMSEu entre les valeurs de référence de rendement, teneur en protéine et reliquat à la récolte et les valeurs obtenues d'après l'IAP ou les résultats de l'assimilation. Pour l'IAP, la RRMSEs est nulle : on obtient en effet une seule valeur de *G*, *P* et *R* car l'IAP est la même pour les 20 jeux de données.

La simulation de *G* à partir des valeurs de variables d'entrée estimées par assimilation est de bonne qualité sauf, à nouveau, dans le cas de la méthode SIMPLEXE pour l'année 1976 : la RRMSE obtenue est de 42.7%, alors qu'elle est de 5.2 et 7.8% respectivement pour les méthodes GLUE et MCIP. Pour l'année 2000 en revanche, les RRMSE sont respectivement de 8.1, 3.6 et 5.4%. Ceci est très largement inférieur à l'erreur relative du modèle en simulation directe (18%, *cf.* Chapitre II). La part de l'erreur systématique est très

faible pour les méthodes GLUE et MCIP en 1976. En 2000 par contre, contrairement à ce qui était le cas avec *LAI* et *QN*, la RRMSEs est du même ordre de grandeur que la RRMSEu (pour la méthode GLUE, respectivement 3.3% et 1.6%), mais reste cependant très limitée. L'amélioration des simulations effectuées grâce à l'assimilation par rapport à celles effectuées d'après l'IAP est très nette, ses erreurs étant de 25.9% et 21.3% pour 1976 et 2000 respectivement. L'effet de l'année utilisée est surtout marqué sur la méthode SIMPLEXE : si les résultats pour cette méthode en 1976 sont mauvais, ils sont en revanche plus proches de ceux des autres méthodes en 2000.

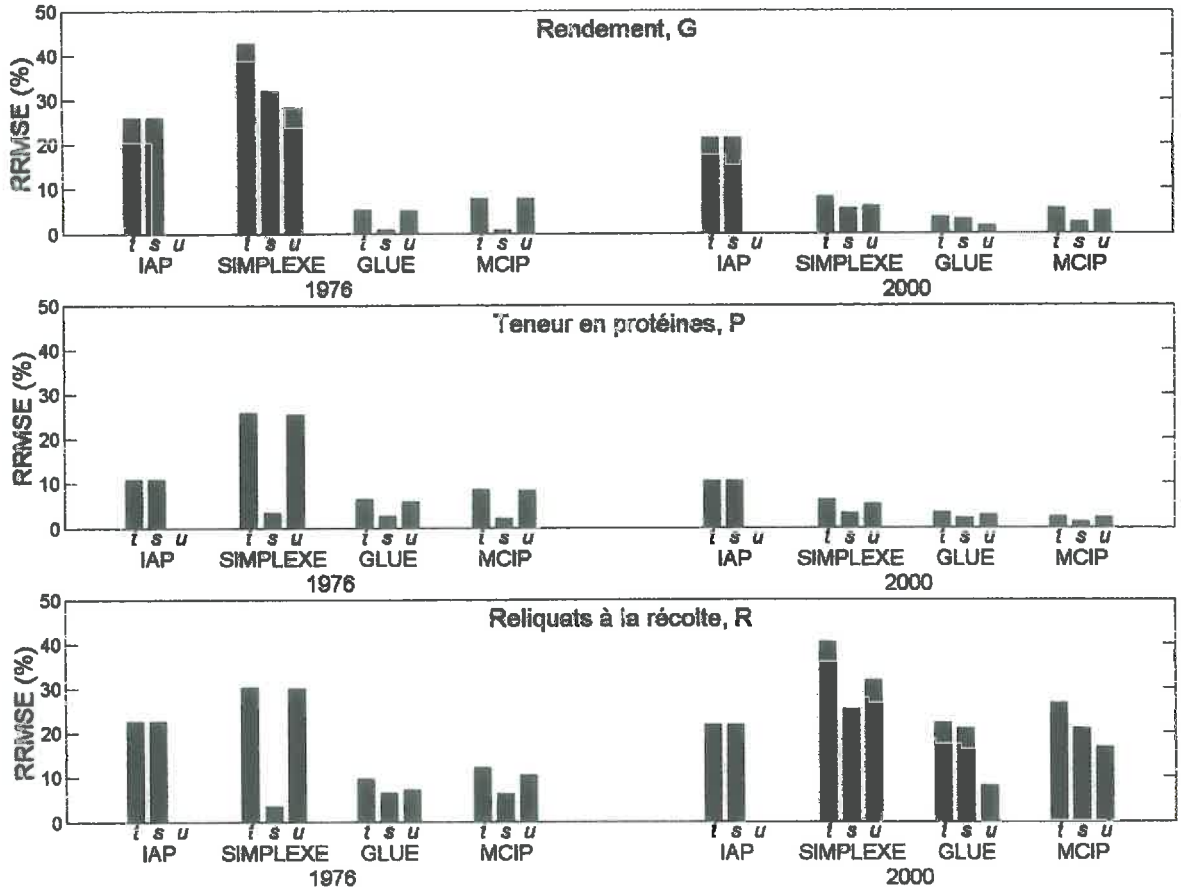


Figure III.3-15 : RRMSE entre les valeurs de référence de *G*, *P* et *R* et les valeurs simulées à partir de l'IAP ou des variables d'entrées estimées (calculées sur les 20 points) pour les années climatiques 1975-1976 (1976 sur la figure) et 1999-2000 (2000). Les valeurs de RRMSE (*t*), RRMSEs (*s*) et RRMSEu (*u*) sont données pour les simulations réalisées d'après l'information *a priori* et d'après les valeurs de variables d'entrée obtenues par assimilation par les méthodes SIMPLEXE, GLUE et MCIP.

Des observations très similaires peuvent être faites à propos de *P* : les RRMSE obtenues sont bien plus faibles que les RRMSE obtenues en simulation directe (21%, *cf.* Chapitre II). Pour les méthodes GLUE et MCIP, elles se situent en dessous de 10% en 1976 et de 5% en 2000. Cependant, l'amélioration des simulations par rapport à l'IAP est nettement

moins marquée, surtout en 1976. Les simulations restent en revanche très peu biaisées comme en témoignent les faibles valeurs de RRMSEs.

En ce qui concerne les résultats sur le reliquat d'azote pour l'année 2000, on constate qu'aucune des trois méthodes ne permet d'améliorer les simulations par rapport à l'IAP : cette dernière a une RRMSE de 21.8%, contre 40.4, 22.1 et 26.5% respectivement pour les méthodes SIMPLEXE, GLUE et MCIP. Cela signifie que l'on n'arrive pas, en assimilant *LAI* et *QN*, à retrouver des valeurs θ_n^{est} qui permettent de simuler correctement *R*. On a vu qu'en 2000, c'étaient surtout les deux paramètres plante, *durvieF* et $vlai_{max}$ qui étaient bien estimés : il semble donc qu'en absence de stress hydrique, le fait d'utiliser des variables caractéristiques de la culture (*LAI* et *QN*) pour l'assimilation ne permette pas de déterminer des variables d'entrée caractéristiques du sol conduisant à des simulations correctes de *R*. Ce constat n'est pas aussi vrai dans les conditions de 1976 : les RRMSE sur *R* obtenues avec les méthodes GLUE et MCIP (respectivement 9.7 et 12.2%) sont assez faibles par rapport à celle obtenue à partir de l'IAP (22.5%). Or, on a vu que pour cette même année, une des variables d'entrée caractéristique du sol était mieux estimée que les autres, *OBSTARAC*. Elle joue particulièrement sur les ressources en eau de la plante et donc, en conditions sèches, sur *LAI* et *QN*. Ceci se traduit par une estimation des variables d'entrée θ_n^{est} qui conduit à une meilleure restitution de *R*. Cela montre l'importance des conditions climatiques sur la qualité d'estimation de *R* lorsque des données caractéristiques du couvert sont utilisées pour réaliser l'assimilation.

III.3.3.3.4 Temps de calcul

D'un point de vue pratique, le temps de calcul, traduit par le nombre de simulations N_S effectuées pour réaliser l'assimilation, est fondamental, malgré les progrès informatiques. En cela, c'est un critère important à considérer pour choisir une méthode avant toute utilisation opérationnelle. Trois facteurs influencent ce nombre de simulations : la méthode d'assimilation utilisée, le nombre *Np* de pixels pour lesquels on doit réaliser une assimilation et le nombre *Ni* de zones différentes de la parcelle sur lesquelles on possède une information *a priori* (dans notre étude, nous avons considéré que N_i=1 car l'IAP était la même pour les 20 jeux de données).

La méthode SIMPLEXE nécessite environ 350 simulations pour converger sur un pixel ; par contre, il est nécessaire d'utiliser plusieurs vecteurs initiaux : nous avons choisi d'en utiliser 25. Le nombre d'IAP sur la parcelle *Ni* n'influence pas N_S. On a donc :

$N_S^{SIMPLEXE}=350\times25\times Np$ Équation III.3-15

La méthode GLUE nécessite $N_G = 200\,000$ simulations par nombre d'IAP sur la parcelle, *Ni*. Cette valeur de N_G sera discutée dans la section III.3.4.1. Le nombre *Np* de pixels n'influence pas N_S. On a donc :

$N_S^{GLUE}=200000\times Ni$ Équation III.3-16

La méthode MCIP nécessite pour l'itération 1 $N_M = 10\,000$ simulations par nombre d'IAP sur la parcelle *Ni*, quel que soit *Np*. Pour les itérations suivantes, les $N_M = 10\,000$ simulations sont à effectuer pour tous les pixels indépendamment du nombre d'IAP sur la parcelle. Pour 5 itérations, on a donc :

$N_S^{MCIP}=10000\times Ni+5\times10000\times Np$ Équation III.3-17

On pourrait se limiter à deux itérations, comme préconisé par Demarty (2001). Ceci donnerait alors :

$N_S^{MCIP*}=10000\times Ni+2\times10000\times Np$ Équation III.3-18

La Figure III.3-16 représente le nombre de simulations nécessaires pour chaque méthode pour des valeurs de *Np*=0,…,250. La valeur N_p=250 correspond à une parcelle de 10 ha et des pixels de 20×20m^2. Quatre hypothèses sont formulées sur la richesse de l'IAP : une seule information pour toute la parcelle (a), trois zones à IAP différentes (b), cinq zones (c) ou sept zones (d). Pour la méthode MCIP, les deux hypothèses sur le nombre d'itérations utilisées sont représentées.

Des trois méthodes testées, il apparaît que la méthode GLUE devient la moins coûteuse en nombre de simulations dès lors que le nombre de pixels commence à être élevé et comparable avec ce que l'on a sur une parcelle de 10 ha, et ce même si le nombre d'IAP est important.

III.3.3.3.5 Conclusion de la comparaison des trois méthodes d'assimilation

Plusieurs points ressortent de cette comparaison de méthodes.

Premièrement, même si les valeurs des variables d'entrée estimées par assimilation, θ_n^{est}, ne correspondent pas à celles qui ont servi à réaliser les simulations de référence, θ_n^{ref}, elles permettent de simuler des valeurs de *LAI*, *QN*, *G* et *P* en assez bon accord avec les

valeurs de référence. Il est cependant nécessaire d'évaluer l'effet de l'emploi de ces variables d'entrée sur la qualité des simulations réalisées pour d'autres années climatiques. Il est

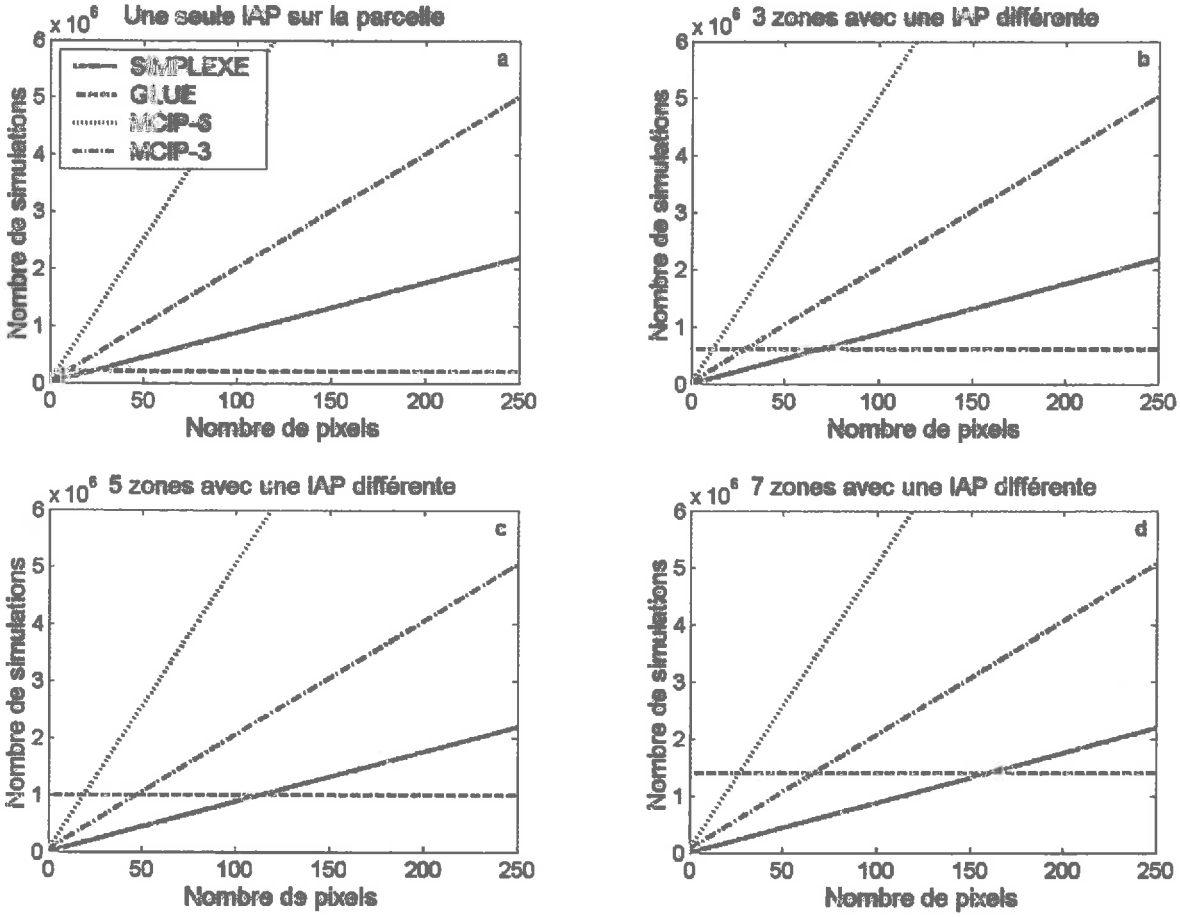


Figure III.3-16 : Nombres de simulations nécessaires pour réaliser des assimilations selon les méthodes SIMPLEXE, GLUE et MCIP en fonction : (i) du nombre de pixels à assimiler (abscisses) et (ii) du nombre d'informations *a priori* différentes sur l'ensemble des pixels (sous-graphiques). MCIP-6 : méthode MCIP appliquée avec 5 itérations ; MCIP-3 : avec 2 itérations.

également utile d'évaluer l'impact des erreurs des observations sur la qualité de l'estimation des variables d'entrée. Ces deux aspects feront l'objet d'une étude dans la partie suivante avec la méthode GLUE.

L'année climatique qui sert à réaliser les assimilations est prépondérante sur la qualité des résultats. On pourrait imaginer d'utiliser plusieurs années contrastées pour réaliser les assimilations afin d'améliorer l'estimation des variables d'entrée. Cet effet des années peut s'expliquer par la différence entre l'intensité des stress hydrique et azoté d'une année sur l'autre, ainsi que par leur positionnement par rapport aux dates d'observation. Ceci pose le

problème de l'échantillonnage temporel des observations qui sera également traité dans la partie suivante.

La variable *R* est particulièrement mal simulée à partir des estimations des variables d'entrée par assimilation de *LAI* et de *QN*. Il est probable que pour être capable d'estimer des valeurs de variables d'entrée qui conduisent à de bonnes simulations de *R*, il faudrait utiliser des variables observées plus proches de *R* pour effectuer l'assimilation. La géophysique pourrait apporter des solutions à ce problème. L'effet de l'année est de ce point de vue particulièrement critique. Ceci conduit à s'interroger sur la nature des variables utilisées pour l'assimilation : on peut d'une part tester l'intérêt d'utiliser *LAI* et *QN* et d'autre part envisager d'assimiler des variables plus caractéristiques de l'état du sol.

Les faibles performances de la méthode SIMPLEXE sont sujets à discussion : il est étonnant que les variables d'entrées et observées soient mal estimées, alors même qu'elles sont toutes prises en compte dans le critère à minimiser. Ceci nécessiterait un approfondissement. On peut essayer de trouver un début d'explication dans la forme du critère minimisé. Si pour une variable d'entrée donnée θ_j, la valeur de l'information *a priori* μ_j^{ve} est éloignée de la valeur de référence θ_j^{ref}, on peut penser que l'algorithme de recherche aura du mal à trouver θ_j^{ref}, car il sera ramené vers μ_j^{ve}. En outre, pour $\theta_j = \theta_j^{ref}$, la valeur du critère n'est pas nulle : on peut donc imaginer que l'algorithme puisse conduire à des valeurs de θ_j donnant une valeur de ce critère plus faible que pour $\theta_j = \theta_j^{ref}$.

Les résultats des méthodes GLUE et MCIP sont comparables en terme de performance d'estimation des valeurs des variables d'entrée du modèle. Un meilleur réglage de tous les facteurs nécessaires à la méthode MCIP pourrait permettre d'améliorer ses performances mais, d'un point de vue pratique, la méthode GLUE est beaucoup plus simple et rapide à mettre en œuvre pour un grand nombre de pixels. En ce qui concerne cette dernière, l'hypothèse de base – l'indépendance et la normalité des erreurs du modèle – n'est pas non plus vérifiée, les valeurs de *LAI* et de *QN* étant fortement corrélées. L'hypothèse d'indépendance pourrait être levée en introduisant des termes de corrélation des erreurs dans le critère, mais cela ferait autant de paramètres à estimer. Malgré cette approximation, elle donne des résultats corrects, mais c'est une limite dont il faut rester conscient.

C'est donc la méthode GLUE qui sera retenue pour réaliser des assimilations à l'aide de jeux de données réelles. La partie suivante permet d'explorer un peu plus l'utilisation de

cette méthode pour l'assimilation de données issues de la télédétection dans un modèle de culture, toujours sur des jeux de données simulées.

III.3.4 Analyse de différentes configurations d'assimilation avec la méthode GLUE

Cette partie est entièrement consacrée à des résultats obtenus avec la méthode GLUE. La plupart des traitements décrits ont aussi été réalisés avec la méthode MCIP mais nous ne ferons que l'évoquer. Une fois le choix de la méthode arrêté sur GLUE, des questions restées en suspens seront abordées :

- le nombre de simulations initiales (N_G) à réaliser avec la méthode GLUE ;
- l'étude de la position et du nombre de dates d'observation sur la qualité des assimilations ;
- l'effet du nombre de variables servant à l'assimilation ;
- l'effet de l'importance des erreurs de mesures sur les résultats de l'assimilation, traduites par le bruit rajouté lors de la création des pseudo-observations. Ceci permettra de déterminer quelques caractéristiques que le jeu de données de télédétection utilisé pour l'assimilation devra posséder pour répondre à nos objectifs ;
- les conséquences de l'utilisation de jeux de données estimés par assimilation lors d'une année climatique donnée sur d'autres années climatiques.

Les données utilisées pour cette partie sont les 20 jeux créés pour la partie précédente.

III.3.4.1 Effet du nombre de simulations initiales dans la méthode GLUE

Pour évaluer l'effet du nombre de simulations initiales N_G sur la qualité de l'assimilation réalisée avec la méthode GLUE, plusieurs séries d'assimilations ont été réalisées avec un nombre variable de simulations prises en compte : 10 000, 20 000, 50 000, 100 000, 200 000, 300 000 et 400 000. Les résultats sur les estimations des variables d'entrée sont donnés sous la forme de RRMSE normalisées par rapport aux RRMSE de l'IAP (sauf dans le cas de $vlai_{max}$) dans la Figure III.3-17.

On aurait pu s'attendre à ce que, pour toutes les variables d'entrée, la qualité de l'estimation augmente avec le nombre de simulations N_G. Ceci est le cas pour des variables

comme *q0*, *HCCF(1)*, *HCCF(2)* et *durvieF* : les valeurs de RRMSE normalisées par rapport à celle de l'IAP prennent par exemple pour *q0* les valeurs de 1.13, 1.05, 0.98, 0.88, 0.85, 0.84 lorsqu'on passe de N_G=10 000 à N_G=400 000, témoin d'une amélioration régulière de la qualité de l'assimilation avec l'augmentation de N_G. Pour d'autres variables, cette amélioration est beaucoup moins régulière : *CALC*, *OBSTARAC* et *DA2*. Pour *CALC* par exemple, la valeur de la RRMSE normalisée pour N_G=10 000 est de 1.15, alors que pour

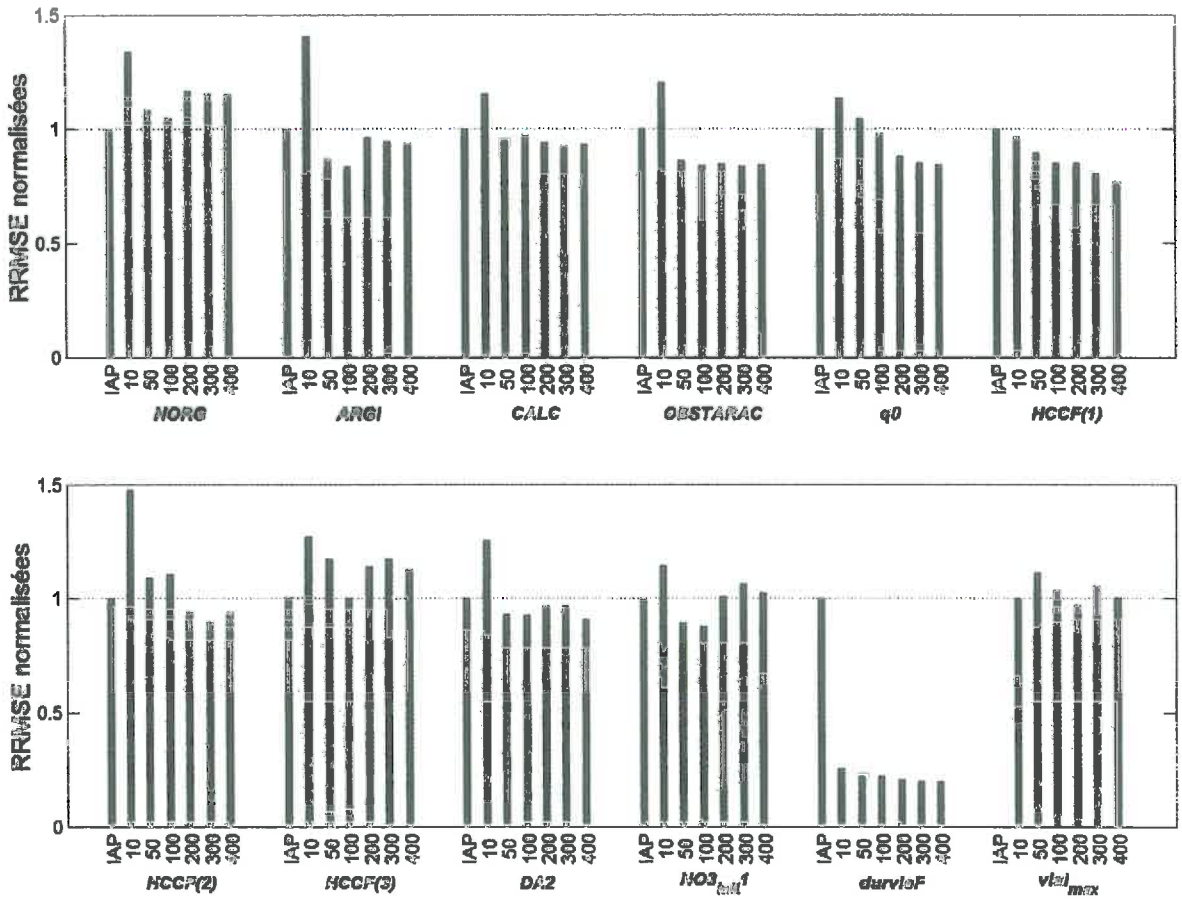


Figure III.3-17 : Effet du nombre de simulations initiales dans la méthode GLUE sur l'estimation des valeurs de chaque variable d'entrée. En abscisse sont indiquées les valeurs du nombre de simulations initiales testées divisées par 1000. Les valeurs de RRMSE sont normalisées entre 0 et 1 par rapport à la valeur obtenue avec l'IAP, sauf dans le cas de $vlai_{max}$ pour laquelle la RRMSE est nulle. La valeur obtenue pour 10 000 a alors été utilisée pour normaliser. L'année climatique utilisée est 2000.

toutes les autres valeurs de N_G, elles se situent autour de 0.95. Pour d'autres variables, comme *NORG*, *ARGI*, *HCCF(3)* et $NO3_{init}1$, 100 000 simulations semblent donner de meilleurs résultats qu'un nombre supérieur. Pour $vlai_{max}$ enfin, l'évolution n'est pas régulière. Il est donc mal aisé de tirer une conclusion, si ce n'est qu'il aurait fallu envisager un plus grand nombre de simulations afin de vérifier si au-delà de 400 000 simulations on a ou non

une stabilisation. Beven & Freer (2001) utilisent 60 000 simulations, Makowski *et al.* (2002), 100 000 ou 200 000. La gamme de nombre de simulations allant jusqu'à 400 000 semble donc pourtant déjà raisonnable.

Si nous examinons les conséquences sur la simulation des variables servant à l'assimilation, nous constatons que N_G a assez peu d'influence sur la qualité de la simulation de *LAI* et de *QN* (Figure III.3-18). En ce qui concerne le rendement, un accroissement du nombre des simulations contribue à diminuer la part non-systématique de l'erreur, mais augmente en revanche la part systématique (RRMSEs=2.5 et RRMSEu=2.5 pour N_G=10 000 et RRMSEs=3.2 et RRMSEu=1.5 pour N_G=400 000). L'erreur totale est à peu près constante. Pour la teneur en protéines, les deux types d'erreur augmentent avec le nombre de simulations. Mais pour le *G* comme pour *P*, les erreurs sont très faibles par rapport à ce qui est obtenu sans assimilation. Enfin, le reliquat présente des résultats plus conformes aux attentes, à savoir que les erreurs diminuent avec le nombre de simulations. Néanmoins, comme on l'a vu, les résultats sont du même ordre de grandeur que ceux obtenus avec l'IAP.

Ces résultats pour le moins surprenants pourraient être liés à une mauvaise utilisation de la méthode ; mais mis à part le nombre de simulations nécessaires, seul le choix de la fonction de vraisemblance peut être remis en cause. Une autre explication consisterait à dire que, par exemple par rapport à ce qu'on observe sur la teneur en protéines, les valeurs de RMSE obtenues pour les différentes valeurs de nombre de simulations se valent et que les différences, somme toute légères, seraient liées à l'échantillonnage. Pour réellement pouvoir trancher cette question, il faudrait donc réitérer cette analyse avec des échantillonnages fort divers.

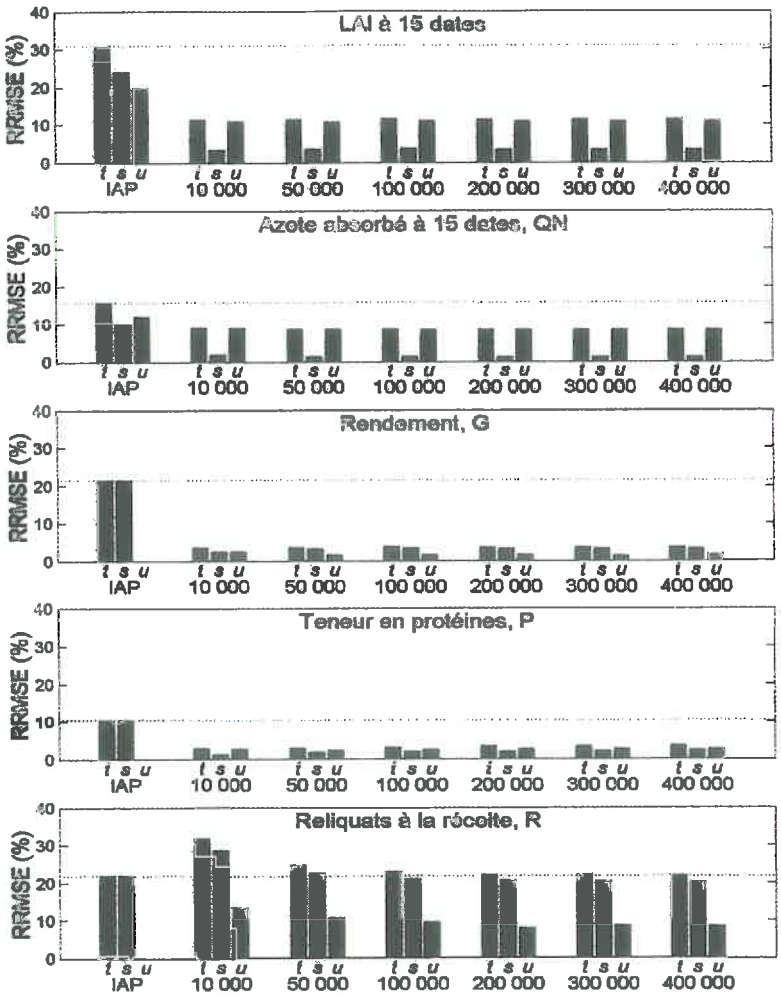


Figure III.3-18 : Effet du nombre de simulations initiales dans la méthode GLUE sur la restitution des variables d'état servant à l'assimilation et des variables d'intérêt final. En abscisse sont indiquées les valeurs du nombre de simulations initiales testées. Les RRMSE (*t*), RRMSEs (*s*) et RRMSEu (*u*) sur ces variables sont indiquées. L'année climatique utilisée est 2000.

III.3.4.2 Effet du nombre de dates de mesure et de leur position dans le cycle

Les images de télédétection qui peuvent servir à estimer *LAI* et *QN* ne sont pas forcément accessibles en grand nombre et sur n'importe quelle partie du cycle. En outre, par rapport à notre travail, pour piloter la fertilisation grâce à l'assimilation dans le modèle de culture, il faudrait utiliser des données acquises avant les apports à moduler, donc avant le mois d'avril. L'objectif de cette partie est d'examiner sur des jeux de données simulées l'effet du nombre de dates d'observation et de leur répartition au cours du cycle cultural sur la qualité de l'assimilation. Les assimilations décrites précédemment comportaient 15 dates

d'observation comme décrit p. 192. D'autres assimilations ont été effectuées qui n'utilisent que 5 de ces dates, réparties de façons diverses :

- tout le long du cycle (01/03/00, 05/04/00, 26/04/00, 17/05/00, 21/06/00), notées '5 rép' ;

- en début de cycle (16/02, 01/03, 15/03, 29/03 et 05/04), notées '5 déb' (ces dates correspondent aux dates de mesures qui pourraient être disponibles pour piloter la fertilisation azotée) ;

- en milieu de cycle (12/04, 19/04, 26/04, 03/05, 10/05), notées '5 mil' (ces dates d'observation ne sont plus utiles pour piloter la fertilisation azotée) ;

- en fin de cycle (17/05, 24/05, 07/06, 21/06, 05/07), notées '5 fin' (même remarque) ;

sur la Figure III.3-19 (ce qui est noté '*LAI* seul' dans cette figure sera commenté dans la partie suivante). Le premier constat est que toutes les configurations d'assimilation permettent d'améliorer les simulations par rapport à l'IAP, sauf dans le cas du reliquat (lignes en pointillé). On constate qu'un nombre important de dates d'observation (15) (lignes en tirets) ne garantit pas une meilleure qualité d'assimilation qu'un nombre plus limité (5) : en effet, la qualité de simulation est du même acabit pour les expériences '15 dates', '5 dates rép' et '5 dates fin'. Les moins bonnes assimilations sont obtenues pour '5 dates déb', qui se situent à un niveau de qualité proche de l'IAP, sauf en ce qui concerne la variable *R*. Ceci suggère que pour mieux estimer celle-ci, il faudrait peut-être utiliser des observations précoces, alors que pour les autres variables, les dates de milieu et de fin de cycle sont plus importantes. Par ailleurs, ces remarques ont des conséquences qui pourraient s'avérer fondamentales pour le reste de notre étude. En effet, si les dates d'observations les plus utiles sont situées tout au long du cycle et plus particulièrement à la fin, vers les mois d'avril à juin, cela réduit la possibilité d'utiliser l'assimilation de données issues de la télédétection dans l'année en cours pour le raisonnement de la fertilisation azotée qui doit se faire vers le mois de mars. Il faudrait plutôt envisager d'utiliser les données acquises au cours de campagnes précédentes afin d'estimer les propriétés permanentes de la parcelle ; les propriétés non permanentes comme le reliquat au semis (par exemple $NO3_{init}1$) devraient être estimées chaque année par mesure ou assimilation des premières données acquises. Nous reviendrons sur cette discussion en temps voulu.

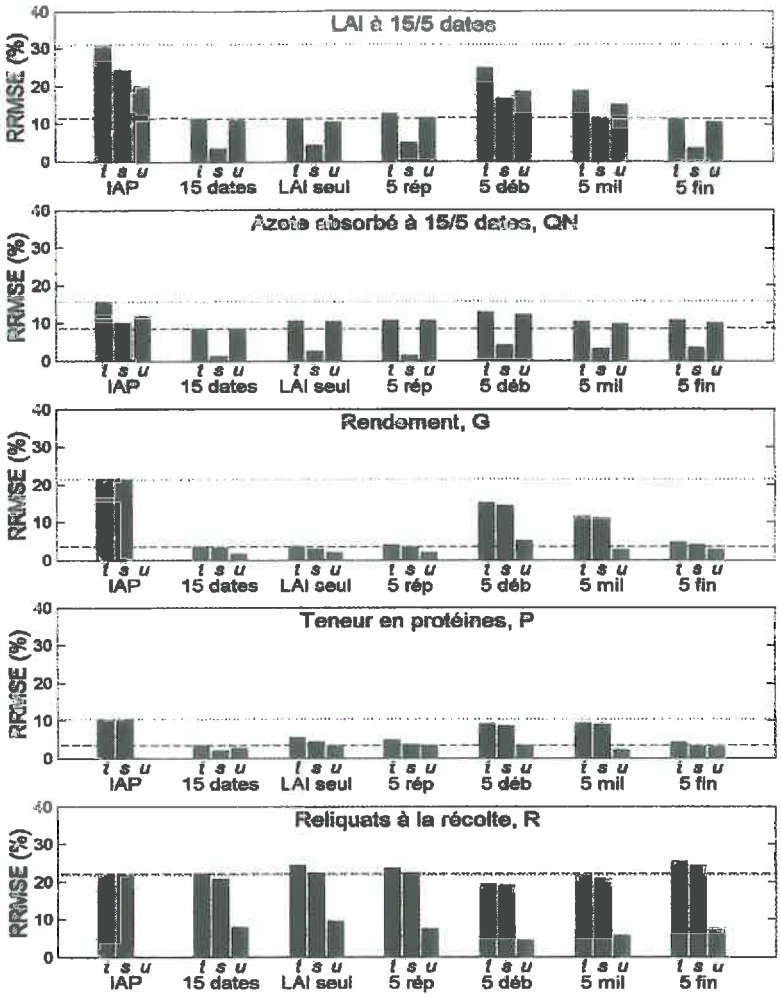


Figure III.3-19 : Effet du nombre de variables et de dates utilisées pour réaliser l'assimilation avec la méthode GLUE sur la restitution des variables d'état servant à l'assimilation et d'intérêt final. '15 dates' : 15 dates d'observations, *LAI* et *QN* utilisées ; '*LAI* seul' : 15 dates, *LAI* utilisé seulement ; '5 rép/déb/mil/fin' : 5 dates d'observations utilisées, respectivement réparties au cours du cycle, placées en début, en milieu et en fin de cycle. Les RRMSE (*t*), RRMSEs (*s*) et RRMSEu (*u*) sur ces variables sont indiquées. L'année climatique utilisée est 2000. La ligne en pointillé indique la RRMSE totale de l'IAP ; la ligne en tirets celle de l'assimilation avec 15 dates et *LAI* et *QN* utilisées.

Une étude menée avec la méthode MCIP mais non présentée ici sur l'effet de la localisation des dates indiquait également que lorsque seules les dates de début de cycle étaient utilisées pour réaliser l'assimilation, les résultats obtenus étaient moins bons.

III.3.4.3 Effet du nombre de variables utilisées pour l'assimilation

Tous les capteurs de télédétection ne permettent pas d'accéder à *QN* via l'estimation du contenu en chlorophylle du couvert. Il est donc utile d'étudier l'intérêt d'utiliser conjointement *LAI* et *QN* pour réaliser l'assimilation. Une assimilation a donc été menée en

n'utilisant comme variables « pseudo-observées » que les différentes dates d'observation du *LAI* et pas celles de *QN*. Les résultats sont présentés dans Figure III.3-19 où l'on compare à présent les résultats de '*LAI* seul' à '15 dates'.

Le fait de n'assimiler que le *LAI* a assez peu de conséquences sur la simulation de cette même variable et du rendement. Par contre, il en a davantage sur les variables qui ont trait à l'azote : *QN* naturellement, la teneur en protéines et le reliquat d'azote. Deux remarques au moins peuvent être faites à ce propos.

D'une part, les valeurs de *LAI* et de *QN* sont fortement corrélées dans le modèle, surtout sur les dates de début et de milieu de cycle, comme illustré dans la Figure III.3-20 qui donne les corrélations de ces deux variables dans le modèle pour 120 000 simulations et pour les 15 dates considérées. Pour les dates de fin de cycle, ceci est moins vrai à cause de la diminution progressive de *LAI*. Il est donc normal, du fait de cette corrélation, que le rajout de *QN* n'améliore que peu l'assimilation car la majorité de l'information contenue dans cette variable l'était déjà dans *LAI*.

D'autre part, ce qui vient d'être dit s'applique surtout pour les variables de type quantité de matière végétale (*LAI* et *G*). Pour celles qui sont plus directement concernées par l'absorption de l'azote, *QN*, *P* et *R*, la dynamique de l'azote présent dans la plante a davantage d'importance sur l'efficacité des simulations de ces variables comme en témoigne la Figure III.3-19. La Figure III.3-21 présente les corrélations entre les trois variables d'intérêt final et le *LAI* et *QN*. Le rendement est fortement corrélé au *LAI* et à *QN* également. Ce peut être pour cette raison que l'assimilation de *LAI* seul suffit à bien le simuler. En revanche, *P* et *R* ne sont pas du tout liés au *LAI*, mais un peu plus à *QN*, quoique très faiblement, ce qui pourrait expliquer le fait que l'assimilation de *QN* apporte de l'information nécessaire pour mieux simuler ces variables.

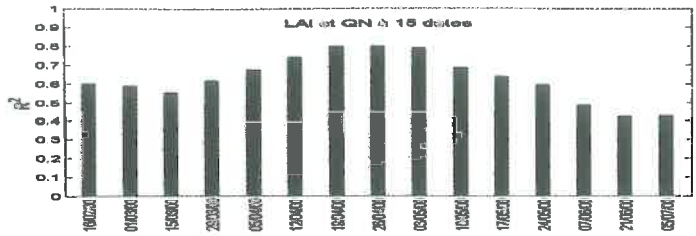


Figure III.3-20 : Représentation sur 120 000 simulations des corrélations entre *LAI* et *QN* pour les 15 dates utilisées dans l'assimilation.

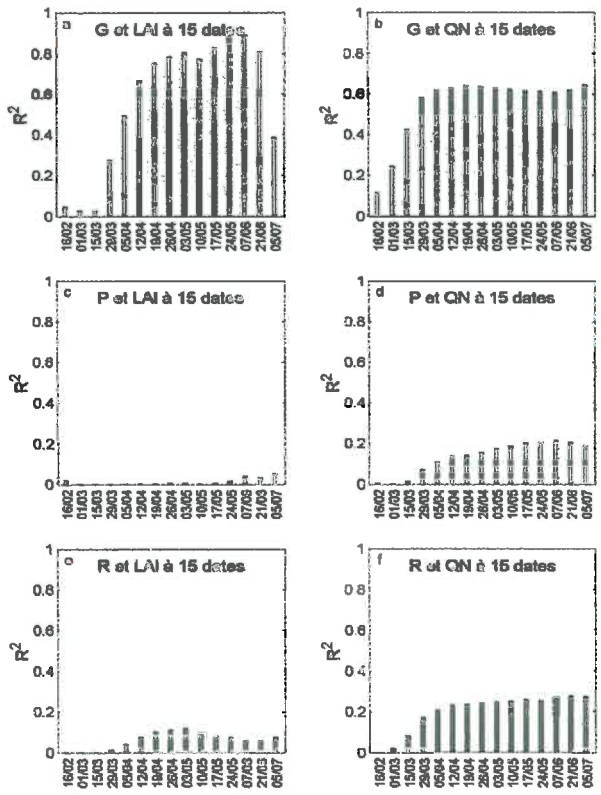


Figure III.3-21 : Représentation sur 120 000 simulations des corrélations entre *G* et *LAI* (a), *G* et *QN* (b), *P* et *LAI* (c), *P* et *QN* (d), *R* et *LAI* (e) et *R* et *QN* (f) pour les 15 dates utilisées dans l'assimilation.

III.3.4.4 Effet de l'importance des erreurs de mesure

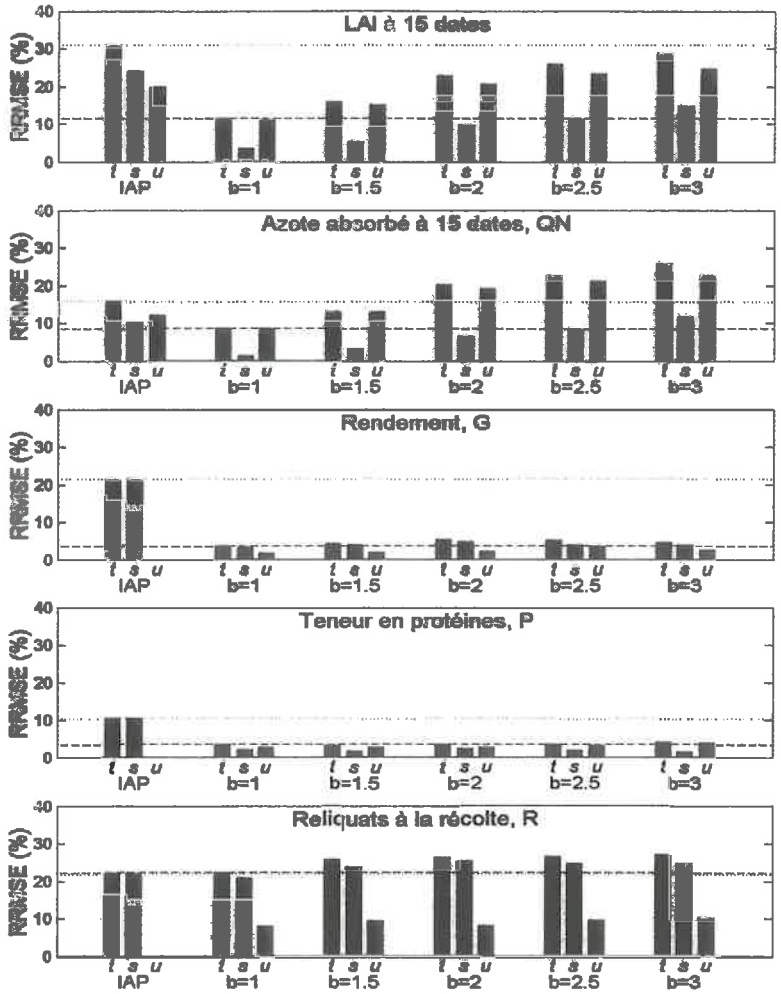


Figure III.3-22 : Effet de l'importance du bruit appliqué aux variables « pseudo-observées » sur la restitution des variables d'état servant à l'assimilation et d'intérêt final après assimilation avec la méthode GLUE. En abscisse sont indiquées les valeurs du coefficient multiplicatif appliqué au bruitage $k \circ Y_n^{po}$. Les RRMSE (*t*), RRMSEs (*s*) et RRMSEu (*u*) sur ces variables sont indiquées. L'année climatique utilisée est 2000. La ligne en pointillés indique la RRMSE totale de l'IAP ; la ligne en tirets celle de l'assimilation avec le bruit minimal.

Dans la méthode de référence, les variables « pseudo-observées » étaient créées en rajoutant au résultat de la simulation par le modèle une variable aléatoire ε_n tirée dans une loi normale $N(0 ; \sigma^2)$ avec $\sigma = k \circ Y_n^{po}$. Afin d'étudier l'importance des erreurs de mesures sur la qualité de l'assimilation, nous nous livrerons dans cette partie à une expérience consistant à faire varier l'ordre de grandeur du bruit, en prenant pour σ la valeur de $\sigma = b.k \circ Y_n^{po}$, b prenant les valeurs successives de 1, 1.5, 2, 2.5 et 3.

Les résultats sont montrés dans la Figure III.3-22. Il est indubitable que l'augmentation de la valeur de *b*, autrement dit du degré de bruitage, a une très grande influence sur la qualité de la simulation après assimilation de *LAI* et de *QN* : au fur et à mesure que le bruitage augmente, la RRMSE totale du *LAI* passe de 11.4% à 15.8, 22.6, 25.8 pour arriver pour b=3 à une valeur de 28.7%, soit pratiquement celle de l'IAP (31%). Pour *QN*, à partir de b=2, l'erreur faite par assimilation est plus grande que celle de l'IAP. Ceci est essentiellement dû à ce que le modèle ne peut pas reproduire une évolution de *LAI* et de *QN* erratique comme c'est le cas des valeurs observées lorsque le bruitage est fort.

Si l'importance des erreurs d'observation a un fort impact sur la qualité de la simulation des variables ayant servi à réaliser l'assimilation, il est irréfragable que les conséquences de l'augmentation du bruit sur la simulation de *G*, *P* et *R* est grandement moins affectée, surtout en ce qui concerne *G* et *P*. En effet, les valeurs de RRMSE de *G* sont respectivement de 3.6, 4.1, 5.2, 5.1 et 4.6% pour des valeurs de b de 1, 1.5, 2, 2.5 et 3, contre 21.3% d'erreur pour l'IAP. Le résultat est similaire pour *P*. La simulation de *R* se dégrade également peu avec l'accroissement de b, mais sa RRMSE dépasse assez largement la valeur obtenue par l'IAP dès b=1.5.

G et *LAI* étant assez bien corrélés, il est cohérent que la dégradation de *LAI* soit concomitante avec une dégradation de *G*. Inversement, la faible dégradation de *P* et *R* peut être liée à leur absence de corrélation avec *LAI* et à leur faible corrélation avec *QN*. Par rapport au fait que la dégradation de *G* est beaucoup plus faible que celle de *LAI* et de *QN*, il est important de remarquer que ce qui est étudié ici est une erreur aléatoire sur la mesure autour de la vraie valeur. Il ne s'agit donc pas d'un biais, la cinétique de *LAI* et de *QN* est conservée. Il est très vraisemblable qu'un bruitage amenant une perturbation systématique des variables observées par l'introduction d'un biais génèrerait des erreurs beaucoup plus importantes sur les variables finales.

L'effet du bruit a également été étudié avec la méthode MCIP ; les résultats ne seront pas développés ici mais les conclusions étaient similaires à celles que nous venons de voir avec la méthode GLUE : si le bruitage affecte la simulation finale des variables bruitées, les répercussions sur les autres variables sont assez limitées.

III.3.4.5 Utilisation des variables d'entrée estimées pour une année climatique donnée sur d'autres années climatiques

Nous avons vu que la méthode GLUE ne permettait pas de retrouver correctement les valeurs des paramètres ayant servi à faire les simulations de référence, θ_n^{ref}. Nous avons également vu que les simulations effectuées avec les valeurs estimées par assimilation, θ_n^{est}, permettaient néanmoins d'obtenir des valeurs de *LAI*, *QN*, *G* et *P* assez proches de celles qui étaient simulées par θ_n^{ref}, ce qui nous a conduit à formuler l'hypothèse que θ_n^{est} était un équivalent de θ_n^{ref}. Nous avons par ailleurs montré que les dates qui permettent de mieux caractériser les variables sol par assimilation appartenaient à une période située après la date de raisonnement de la fertilisation azotée, ce qui implique que des valeurs de θ_n^{est} estimées une année donnée devraient être utilisées pour faire des préconisations une autre année. Le but de cette partie est donc de vérifier que θ_n^{est} est bien un équivalent de θ_n^{ref}, y compris pour d'autres années climatiques.

Nous avons donc considéré 20 années climatiques différentes des années 1975-1976 et 1999-2000. Nous avons fait des simulations pour ces 20 années climatiques et les 20 jeux de données avec les valeurs des variables d'entrée de référence θ_n^{ref} et avec les valeurs θ_n^{est} estimées sur l'année climatique 1975-1976 d'une part et 1999-2000 d'autre part ainsi qu'avec l'IAP. La Figure III.3-23 compare les résultats de ces simulations en terme de RRMSE entre les valeurs simulées avec θ_n^{ref} et les valeurs simulées avec des θ_n^{est} ou l'IAP.

L'utilisation des valeurs de variables d'entrées estimées par assimilation sur les séries climatiques 1976 et 2000 permet d'obtenir des valeurs de *LAI* et *QN* de bien meilleure facture que celles obtenues avec l'IAP pour la série de 20 années climatiques : alors que l'IAP permet d'obtenir une RRMSE de 28.6% sur le *LAI* , le jeu de valeurs de variables d'entrée estimé par la méthode GLUE pour l'année climatique 1975-1976 donne 8.5% et 1999-2000, 13%. Pour *QN*, on a respectivement 15.7%, 5.3% et 6.9%. Ces simulations sont par ailleurs très peu biaisées. On constate que les résultats sont meilleurs avec le jeu obtenu pour l'année 1976, c'est-à-dire l'année où les facteurs de stress s'expriment le plus, ce qui confirme l'importance de l'année climatique sur la qualité de l'assimilation.

L'effet de l'année d'assimilation est encore plus marqué sur les variables d'intérêt final. Alors que l'IAP permet d'obtenir une RRMSE de 21.4% sur le rendement, le jeu de données estimé sur l'année 1976 donne 5.1% et 2000, 10.8% (la part de l'erreur systématique

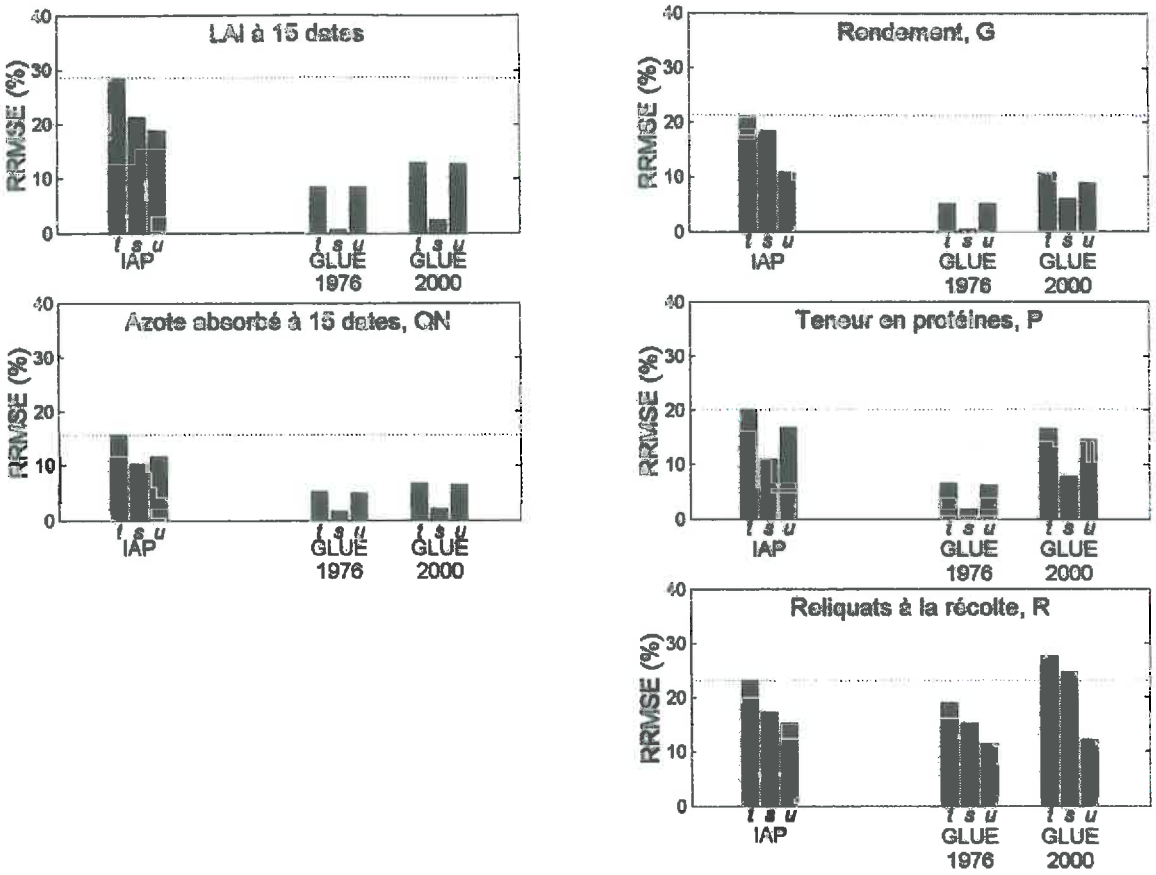


Figure III.3-23 : Sur 20 séries climatiques, représentation pour les variables d'état *LAI*, *QN*, *G*, *P* et *R* des RRMSE entre les valeurs simulées avec θ_n^{ref} et avec θ_n^{est}. Les valeurs de RRMSE (*t*), RRMSEs (*s*) et RRMSEu (*u*) sont données pour les simulations réalisées d'après l'information *a priori* et d'après les valeurs de variables d'entrée obtenues par assimilation par la méthode GLUE ; 1976 : valeurs retrouvées en utilisant l'année climatique 1975-1976 pour réaliser les assimilations ; 2000 : en utilisant l'année climatique 1999-2000.

étant beaucoup plus marqué que pour 1976), ce qui corrobore ce qui vient d'être dit sur l'importance de l'année climatique utilisée pour l'assimilation. Ceci est encore plus marqué pour la teneur en protéines pour laquelle le jeu estimé sur l'année 2000 améliore peu les simulations sur 20 ans par rapport à l'IAP. Pour *R*, l'effet année est également marqué : les résultats de 1976 donnent des simulations de qualité légèrement meilleure à celle de l'IAP, alors que ceux de 2000 en donnent des pires, ce qui confirme la difficulté de déterminer des valeurs de variables d'entrée qui ont un sens du point de vue de *R* en assimilant *LAI* et *QN*, surtout une année humide.

Il résulte de ce qui précède que :

- bien que les valeurs de variables d'entrée estimées ne soient pas correctes, elles permettent de réaliser des simulations sur d'autres années climatiques qui donnent des résultats acceptables ;

- la qualité des estimations permise dépend de la série climatique qui a servi à effectuer l'assimilation : une année humide (2000) ne permettra pas d'estimer correctement des valeurs qui donneront des simulations correctes sur d'autres années climatiques ; une année sèche (1976) permet d'extérioriser le rôle des propriétés du sol sur la croissance de la culture et donc de déterminer des valeurs de variables d'entrée sol qui servent à chaque fois que le facteur limitant lié au sol joue.

Cette étude menée avec la méthode MCIP a montré des résultats très similaires.

III.3.5 Discussion et conclusions

Ce qui suit se veut être une synthèse des points essentiels qui ont été abordés au cours de cette partie et donne quelques éléments de réflexion pour la suite.

III.3.5.1 Simulation des variables d'intérêt final par assimilation

Si deux des méthodes testées, GLUE et MCIP, permettent de déterminer des vecteurs de variables d'entrée qui conduisent à des simulations du rendement et de la teneur en protéines significativement meilleures que celles qui sont obtenues à partir d'une connaissance fruste de la parcelle (IAP), l'estimation du reliquat d'azote est en revanche beaucoup plus problématique, et ce sans inclure les erreurs du modèle. Si certains types d'années climatiques (sèches) permettent d'estimer des propriétés du sol qui reproduisent assez bien le reliquat, d'autres (humides) ne le permettent pas du tout.

Par rapport à la problématique du raisonnement de la fertilisation azotée utilisant un critère agro-environnemental qui prend en compte une variable environnementale, les résultats du Chapitre II et de celui-ci convergent pour montrer la difficulté d'utiliser le reliquat d'azote à cette fin : (i) ils sont mal simulés par le modèle ; (ii) sur des jeux de données expérimentaux ils présentent de moins bonnes propriétés pour le raisonnement de la fertilisation que la variable bilan environnemental (résultats du Chapitre II) ; (iii) en assimilant *LAI* et *QN* dans un modèle de culture, on ne peut pas prétendre réduire suffisamment l'incertitude sur l'estimation du reliquat d'azote. En revanche, comme le rendement et la teneur en protéines sont assez bien retrouvés, il en ira de même du bilan environnemental car cette dernière est déterminée en fonction des deux premières. C'est pourquoi nous n'avons volontairement pas inclus cette dernière variable dans les résultats de cette partie afin de ne pas trop les alourdir. Dans le chapitre suivant qui aboutira à la

préconisation d'une carte de fertilisation azotée, nous utiliserons donc le bilan azoté et non le reliquat comme variable environnementale.

Par ailleurs, nous avons vu que les dates d'observation qui permettent le mieux de réaliser l'assimilation sont celles qui sont situées plutôt en milieu, voire en fin de cycle. Or, le raisonnement de la fertilisation azotée se fait assez tôt. Ce simple constat contrecarre fortement le dessein initial de la méthode qui était d'utiliser les données de télédétection de l'année *n* pour raisonner la fertilisation azotée de cette même année et suggère qu'il sera nécessaire d'utiliser les données de l'année *n*-1, voire de plusieurs années antérieures. Pour les propriétés non permanentes de la parcelle comme la variable *NO3initf(1)* dont on a vu l'importance, ceci est fâcheux car il faudrait vraiment pouvoir les estimer pour l'année en cours. En outre, même si on a vérifié sur des données virtuelles la compatibilité de jeux de variables d'entrée sur des années climatiques autres que celles qui ont servi à les estimer, il sera important de vérifier sur des données expérimentales que des variables estimées une année sont également valables pour d'autres années.

III.3.5.2 Conséquences prévisibles de l'effet des erreurs du modèle

Dans toute cette partie méthodologique, le modèle était considéré comme parfait puisque les observations étaient générées par simulation : le bruitage permettait de représenter les erreurs d'observations. On a vu qu'aucune des méthodes d'assimilation testées ne permettait de bien retrouver les valeurs des variables d'entrée qui avaient servi à créer les simulations de référence.

Dans le cas de l'utilisation de données vraies, issues des mesures de télédétection, on introduit dans le système les vraies erreurs d'observation d'une part et les erreurs du modèle d'autre part, ces dernières ayant été quantifiées dans le Chapitre II. Dans ce cadre, nous ne pourrons plus nous référer à des « vraies » valeurs de variables d'entrée puisque d'une part elles nous sont inconnues et que d'autre part, elles dépendent à la fois du système décrit (le champ) mais aussi de l'outil de description du système (le modèle), qui ne le décrit pas parfaitement (autrement dit, les « vraies » valeurs, celles qui caractérisent le système, ne sont pas forcément les valeurs qui permettent de mieux simuler les grandeurs observées car la description du système, le passage entre les valeurs d'entrée et les valeurs observée, n'est pas parfaite). Il faudra donc se contenter de juger de l'efficacité de l'assimilation des données issues de la télédétection dans le système en comparant, comme on l'a fait dans cette partie,

les simulations réalisées sans assimilation (uniquement grâce à l'information *a priori*) à celles réalisées avec les valeurs obtenues grâce à l'assimilation.

Comme on a vu que l'effet des erreurs –du bruitage– avec les simulations de référence était important, il y a fort à parier que le cumul des erreurs du modèle et des observations conduise à une description des variables d'intérêt par le modèle assez perturbée par rapport aux observations. Il sera important de parvenir à quantifier l'effet de ces erreurs sur la carte de préconisation.

III.3.5.3 Quelques mots sur les méthodes GLUE et MCIP

Les méthodes GLUE et MCIP ont été développées dans le même contexte (GLUE provient de l'hydrologie, MCIP est dérivée de méthodes également issues de ce domaine d'étude) et répondent aux mêmes objectifs, évaluer une distribution de variables d'entrée qui permettent de reproduire par simulations des valeurs observées de variables de sortie de modèles relativement complexes. Nous venons de comparer ces deux méthodes du strict point de vue de leurs performances dans un cadre bien précis. Il serait intéressant de pousser plus loin la comparaison, notamment en analysant leur façon de déterminer les distributions *a posteriori*. Cette comparaison n'est pas dans le propos de notre travail, mais voici succinctement brossées deux idées d'investigation.

La méthode MCIP réduit l'intervalle de variation des variables d'entrée tout en déterminant celles auxquelles la qualité des simulations par rapport à des observations est la plus sensible. La méthode GLUE approche une distribution *a posteriori* dont on peut connaître la moyenne et l'écart type. Si ce dernier est grand, la distribution *a posteriori* sera proche de la distribution *a priori* et ce sera le signe que la valeur de la variable d'entrée concernée joue peu sur la qualité de la simulation. Au contraire, si l'écart type est faible, cela signifiera que la moyenne trouvée pour cette variable d'entrée permet d'obtenir de bonnes simulations et que s'en éloigner compromet la qualité des simulations. En ce sens, la méthode GLUE réalise de façon implicite une analyse de sensibilité au sens de la méthode MCIP. Il serait donc intéressant de vérifier, pour chaque variable d'entrée, s'il existe une relation entre les valeurs déterminées par la méthode MCIP de P_{KS} et les écarts types trouvés par la méthode GLUE.

Par analogie aux itérations effectuées avec la méthode MCIP, on pourrait envisager d'utiliser la distribution *a posteriori* déterminée par la méthode GLUE pour refaire une phase

de simulation et d'analyse. Ceci pourrait peut-être permettre d'affiner l'estimation de chaque variable d'entrée, en resserrant peu à peu les écarts types trouvés. Mais l'intérêt principal de la méthode GLUE dans la perspective d'une utilisation sur un grand nombre de points, à savoir que toutes les simulations peuvent être utilisées pour tous les points ayant la même information *a priori*, serait perdue : l'information *a priori* des itérations, correspondant à la distribution *a posteriori* de l'itération précédente, serait différente pour tous les points.

L'idée que la méthode GLUE réalise de façon implicite un choix des variables d'entrée à estimer ou non est importante pour la suite de l'étude : cela permettrait de s'affranchir de l'analyse de sensibilité.

III.4 Conclusion

Les deux méthodologies qui viennent d'être développées dans ce Chapitre visent à se doter d'outils capables de renseigner, à l'échelle intra-parcellaire, chacune des variables d'entrée du modèle STICS. Elles sont basées sur des approches différentes.

L'approche pédologique consiste à établir une carte des sols à une échelle compatible avec les impératifs de l'agriculture de précision et à établir un ensemble de règles de pédotransfert qui permettent de renseigner, à partir de cette carte, les différentes entrées sol du modèle. L'effort de caractérisation réalisé n'étant pas envisageable pour un agriculteur, cette approche servira de référence. Néanmoins, une caractérisation de ce type et de moindre ampleur pourrait servir d'information *a priori* pour l'approche suivante.

L'approche par assimilation utilise des données observées par télédétection en cours de culture afin d'estimer spatialement les variables d'entrée du modèle sur lesquelles on possède une information *a priori* rudimentaire. Cette dernière peut être utilisée dans des applications réelles et elle doit donc être comparée à la première en terme de performances.

Ces deux approches présentent chacune des sources d'erreurs incompressibles :

- dans l'approche pédologique, elles proviennent essentiellement de la caractérisation des horizons pédologiques sur la parcelle (cartographie) et des outils permettant le passage entre cette caractérisation et les besoins du modèle (règles de pédotransfert) ;

- dans l'approche par assimilation de données, elles sont générées par les méthodes utilisées, dont on vient de voir qu'elles ne permettent pas de retrouver les variables d'entrée ayant permis de créer des simulations de référence.

A ces erreurs, il faut rajouter celles qui proviennent du modèle, et qui n'ont pas été caractérisées ici. Dans l'approche par assimilation, les erreurs d'estimation des valeurs des variables biophysiques servant à l'assimilation seront également à prendre en considération.

Un des objectifs du chapitre suivant sera de comparer ces deux approches sur le jeu de données expérimentales de Chambry. Comme les mesures effectuées n'ont pas été réalisées aux mêmes échelles, les modalités de cette comparaison devront être discutées.

IV - Quatrième chapitre Mise en œuvre de l'outil de préconisation spatiale de la fertilisation azotée

On a cultivé le désespoir, on en a fait de la littérature, des œuvres d'art.

Ionesco, Le Solitaire

Table des matières

IV.1 Introduction 225
IV.2 Mesures spatialisées de variables caractéristiques de la culture 226
IV.2.1 Etablissement de cartes de rendement et de teneur en protéines 226
IV.2.2 Détermination de variables biophysiques caractéristiques du couvert à partir des mesures de télédétection 228
IV.2.3 Relations entre les données issues de la télédétection et les variables simulées par un modèle de culture 231
IV.2.4 Visualisation des cartes établies à Chambry 237
IV.2.5 Conclusion 244
IV.3 Comparaison de l'approche par cartographie et de l'approche par assimilation 245
IV.3.1 Mise en œuvre de la comparaison 245
IV.3.2 Comparaison des simulations du *LAI* et de *QN* en 2000 et 2001 247
IV.3.3 Comparaison des variables d'entrée 250
IV.3.4 Comparaison des variables de sortie en 2000 et 2001 254
IV.3.5 Comparaison des variables de sortie en 2002 et 2003 260
IV.3.6 Evaluation de la qualité de la simulation de la variabilité spatiale 262
IV.3.7 Discussion sur les assimilations de 2000 et 2001 267
IV.4 Mise en œuvre expérimentale de la méthode 272
IV.4.1 Caractérisation du sol utilisée 272
IV.4.2 Réalisation des simulations croisant scénarios techniques et climatiques 275
IV.4.3 Choix et application du critère agro-environnemental 275
IV.4.4 Passage de la carte préconisée à la carte réellement appliquée 276
IV.4.5 Evaluation expérimentale rapide de l'intérêt de moduler la fertilisation azotée à l'échelle de la parcelle 278
IV.5 Conclusion 280

IV.1 Introduction

Ce dernier chapitre de résultats est la synthèse des deux précédents. Le Chapitre II traitait d'une méthode du choix de la dose optimale d'azote par STICS. Le Chapitre III décrivait la mise en place de deux méthodes de spatialisation des simulations du modèle : l'approche par cartographie – établissement de couvertures pédologiques associées à des règles de pédotransfert et mesure des variables initiales du modèle – et l'approche par assimilation – estimation des variables d'entrée du modèle décrivant le sol par assimilation de données issues de la télédétection. Chacun de ces aspects sera mis en œuvre et évalué sur le jeu de données de Chambry.

Cette évaluation comporte deux volets. Le premier compare les deux approches de spatialisation sur leur aptitude à restituer la variabilité spatiale observée, notamment en terme de rendement. La complémentarité possible entre ces deux approches aurait consisté à réaliser des assimilations sur la base des estimations des variables d'entrée par l'approche cartographique. Elle ne sera pas explorée ici, considérant que la conjugaison de ces deux sources d'information n'est pas envisageable au niveau d'une exploitation agricole. Néanmoins, l'information *a priori* nécessaire à l'utilisation de l'approche par assimilation correspondra au minimum et au maximum des mesures effectuées dans l'approche par cartographie. Le second volet décrit la mise en œuvre de la méthode de préconisation de la fertilisation grâce à la spatialisation du modèle. Ce travail a été effectué dans le cadre de la campagne 2002-2003 et visait à tester en conditions réelles l'intérêt de moduler la fertilisation. A l'époque de sa réalisation, la caractérisation de la variabilité de la parcelle grâce à l'approche par assimilation n'était pas encore prête, ce qui fait que la préconisation a été réalisée sur la base de simulations effectuées grâce à l'approche cartographique.

Avant de décrire ces deux phases, nous décrirons les méthodes d'acquisition de cartes des variables caractéristiques de la culture impliquées dans la spatialisation du modèle et sa validation : les cartes de *LAI* et de quantité d'azote absorbé par la culture (*QN*) sont issues de mesures de télédétection, les cartes de rendement proviennent de l'utilisation d'un capteur de rendement couplé à un SIG et les cartes de teneur en protéines sont établies en krigeant des mesures effectuées ponctuellement.

IV.2 Mesures spatialisées de variables caractéristiques de la culture

Différentes variables caractéristiques de l'état de la culture peuvent être estimées à l'échelle intra-parcellaire. L'utilisation d'un capteur de rendement sur la moissonneuse-batteuse couplé à un Système de Positionnement Géographique (GPS) permet de dresser des cartes de rendement. Des capteurs de teneur en protéines commencent également à apparaître mais ne sont pas pour l'heure opérationnels. Les mesures de réflectance du couvert par télédétection permettent d'accéder au *LAI*, à la teneur en chlorophylle du couvert (*Cab*), à son contenu intégré en chlorophylle (*QCab*, le produit des deux variables précédentes) en utilisant des modèles de transfert radiatif, et à l'azote absorbé par la culture (*QN*) en liant cette dernière à *QCab*. Cette partie présente les méthodes d'établissement des cartes de *LAI* et *QN* et décrit succinctement leur mise en œuvre au cours des expérimentations de Chambry. Les résultats de cette caractérisation seront présentés dans le dernier paragraphe.

IV.2.1 Etablissement de cartes de rendement et de teneur en protéines

La partie bibliographique qui introduit ce passage ne concerne que l'établissement de cartes de rendement.

IV.2.1.1 Aspects de bibliographie

Si on trouve parfois affirmé que l'établissement de cartes de rendement ne pose plus aucun problème technique, il n'en reste pas moins que différentes sources d'imprécision subsistent. Dans la plupart des capteurs de rendement, comme celui qui est utilisé pour cette étude, la quantité de blé est évaluée à travers son volume qui est ensuite converti en masse. Ceci peut être une des principales sources d'erreur (Blackmore & Moore, 1999). La largeur de coupe est supposée constante dans la plupart des logiciels ; or, selon la trajectoire de la moissonneuse-batteuse, la largeur réelle peut varier (Stafford *et al.*, 1997). Le délai entre la coupe et la mesure, accentué par les problèmes occasionnés par l'engorgement aléatoire des grains dans la machine, peut perturber les mesures et leur représentation spatiale (Birrell *et al.*, 1996) et doivent donc être pris en compte et peuvent être modélisés (Maertens *et al.*, 2001). Des erreurs de positionnement dues au GPS peuvent également perturber la représentation cartographique du rendement (Blackmore & Moore, 1999). Enfin, une carte de rendement représente le « rendement machine », inférieur au rendement biologique simulé par

les modèles de culture en raison des pertes au moment du battage (Blackmore & Moore, 1999).

Si des tentatives pour proposer des interprétations visuelles de cartes de rendement d'une ou plusieurs années ont été faites (Doerge, 1999), Paz *et al.* (1999) constatent que les efforts pour corréler les rendements au type de sol, son orientation, sa fertilité ou autres ont été voués à des échecs relatifs. La variabilité temporelle du rendement en est la cause. Comme il est le résultat d'interactions entre le climat, le sol, la plante et les techniques culturales, les zones de faibles rendements mises en évidence une année donnée peuvent devenir des zones à forts rendements dans d'autres conditions ; encore plus gênant, les zones peuvent ne pas être du tout stables, montrant bien que les limites de sol sont loin d'être nettes et constantes dans le temps (Engel, 1997, Doerge, 1999, Paz *et al.*, 1999). A l'exercice de la comparaison de cartes, Wendroth *et al.* (2003) trouvent que le rendement est mieux expliqué par les images de télédétection de la culture que par les propriétés de sol. L'utilisation de modèles de culture permet justement de prendre en compte ces interactions complexes : c'est ce qui amène Basso *et al.* (2001) à proposer d'utiliser les cartes de rendement comme élément de validation des modèles de culture spatialisés à l'échelle de la parcelle.

Si l'établissement des cartes de rendement n'est pas exempt de toute difficulté, il est une pratique courante en agriculture de précision et historiquement à son origine. Pour cette raison notamment, c'est une information majeure à prendre en compte, comme révélateur des interactions sol-plante pouvant aider à mieux les caractériser et comme variable d'intérêt à prédire. Comme l'ont fait par ailleurs Basso *et al.* (2001), nous l'utiliserons principalement comme source de données de validation.

IV.2.1.2 Cartes de rendement et de teneur en protéines établies à Chambry

Des cartes de rendement ont été établies à Chambry en utilisant un système de cartographie du rendement RDS et un GPS Omnistar. La largeur de la barre de coupe était de 5 m. La conjugaison de la fréquence de mesure de volumes de grain et de la vitesse d'avancement de la machine permet d'obtenir une résolution spatiale dans le sens de l'avancement de 6 m environ. La « résolution » de la carte est donc d'environ 6×5 m^2. Afin de valider les mesures effectuées par ce capteur et pouvoir corriger la carte en conséquence, des prélèvements d'échantillons de blé ont été réalisés sur la parcelle au niveau des nœuds des grilles et analysés en laboratoire en terme de rendement en grain et en paille et de qualité des grains. Cette confrontation a permis de montrer qu'il existait une assez faible différence entre

le rendement mesuré par la moissonneuse-batteuse et son capteur et le rendement biologique mesuré sur les échantillons prélevés (Machet *et al.*, 2001, Gomez, 2002). Les cartes de rendement ont été traitées par l'INRA d'Orléans (A. Couturier) pour normaliser les valeurs de rendement à 0% d'humidité, gérer le décalage entre le moment de la coupe et de la mesure du grain et éliminer les points aberrants. Elles ont été établies à l'échelle de pixels de 20×20 m^2 à l'INRA d'Avignon (P. Clastre, M. Pringle).

En outre, des cartes de teneur en protéines ont également été dressées grâce au prélèvement d'échantillons de grains dans la trémie au niveau des points des grilles et sur des diagonales. Les mesures ont permis d'établir par krigeage des cartes de teneurs en protéines sur les pixels de 20×20 m^2 (P. Clastre, M. Pringle). Des cartes de rendement et de teneurs en protéines ont été réalisées de la sorte de 2000 à 2003. Elles seront présentées et discutées dans la partie IV.2.4.

IV.2.2 Détermination de variables biophysiques caractéristiques du couvert à partir des mesures de télédétection

La vocation de cette partie est de brosser succinctement la façon dont les variables biophysiques – indice de surface foliaire (*LAI*), teneur en chlorophylle (*Cab*) et contenu intégré en chlorophylle du couvert ($QCab = LAI \times Cab$) – sont déterminées grâce aux mesures de réflectance par inversion de modèles de transfert radiatif.

IV.2.2.1 Aspects de bibliographie

Les premières tentatives d'estimation des variables biophysiques des couverts végétaux à partir des mesures de réflectance fournies par la télédétection consistaient à exprimer directement les premières en fonction des secondes (*e.g.* Yoder *et al.*, 1995). Des indices de végétation, combinaison entre plusieurs réflectances, étaient également utilisés (*e.g.* Gitelson & Merzlyak, 1997). Ces approches donnent des résultats intéressants, montrant la corrélation entre réflectances ou indices de végétation et variables biophysiques, mais leur caractère fondamentalement empirique les rend malaisément généralisables et utilisables spatialement.

Pour cette raison, des modèles physiques décrivant les relations entre caractéristiques du couvert et réflectance ont été développés. Le modèle PROSPECT (Jacquemoud & Baret, 1990) simule par exemple la réflectance d'une feuille en utilisant comme variables d'entrées ses propriétés biochimiques (dont *Cab*) et structurales. Le modèle SAIL (Verhoef, 1984)

calcule la réflectance du couvert à partir des sorties de PROSPECT, de la structure du couvert (dont *LAI*), de la réflectance du sol et de la configuration géométrique de la mesure. Le couplage de ces deux modèles permet donc d'évaluer la réflectance du couvert à partir de *Cab*, de *LAI* et d'autres variables d'entrée. Ainsi, le principe de détermination des variables *Cab* et *LAI* d'un couvert dont on a mesuré les réflectances pour différentes longueurs d'onde consiste à inverser ces modèles, c'est-à-dire à retrouver les entrées à partir des sorties, ce qui s'apparente fortement aux méthodes d'assimilation que nous avons utilisées. Nous nous bornerons ici à dresser un bref historique de l'inversion du modèle couplé SAIL+PROSPECT utilisé dans le cadre du projet Agriculture de Précision de Chambry.

Jacquemoud *et al.* (1995) réalisent une inversion de ce modèle en utilisant l'algorithme quasi-Newton et en utilisant comme fonction de coût, c'est-à-dire comme critère à minimiser, une MSE (Mean Square Error). Ces auteurs réalisent des inversions avec plus ou moins de variables d'entrée à estimer, les autres étant fixées à leur valeur nominale. Ils constatent que moins le nombre de variables d'entrée à estimer est important, plus leur valeur est correctement retrouvée. Combal *et al.* (2002), pour améliorer cette méthode, partent du constat que l'inversion d'un modèle est par essence un « problème mal posé » (ill-posed problem), à savoir qu'il n'existe pas *a priori* de solution unique et que les erreurs de mesures et les incertitudes liées à la structure du modèle peuvent induire de larges variations des solutions du problème inverse. Les algorithmes de recherche peuvent s'arrêter à des valeurs qui correspondent à des minimums locaux. Un moyen de contourner, au moins partiellement, ce problème, est de restreindre l'espace de recherche des valeurs des variables d'entrée. Une façon de faire, employée par Jacquemoud *et al.* (1995), est de fixer des bornes à chaque variable d'entrée, bornes qui ne peuvent être franchies par l'algorithme de recherche. Une autre solution, étudiée par Combal *et al.* (2002), est d'inclure de l'information *a priori* dans la fonction coût. Selon l'algorithme de recherche utilisé, cette information n'est pas utilisée de la même façon. Pour un algorithme de recherche de type quasi-Newton, ces auteurs proposent d'écrire la fonction coût selon un formalisme proche de celui utilisé dans le Chapitre III (paragraphe 3.2.1.2). Cette utilisation d'une pré-connaissance de la valeur des variables d'entrée permet d'améliorer l'estimation des paramètres. Enfin, Moulin & Guérif (soumis) utilisent cette technique en étudiant l'impact de l'erreur des observations et de l'information *a priori* sur l'estimation de *Cab* et de *LAI*. C'est précisément cette méthodologie qui a été mise en œuvre pour déterminer les variables biophysiques caractéristiques du couvert à partir des mesures de télédétection dans le cadre de l'expérimentation de Chambry. Il convient de

remarquer à ce stade que l'utilisation de l'information *a priori* revient à réduire l'espace de recherche des paramètres autour de valeurs que l'on sait être plausibles. Mais la possibilité de présence de minimums locaux dans cet espace demeure néanmoins, même si elle est plus faible.

Les trois auteurs cités ci-dessus constatent que le produit de *LAI* et de *Cab*, *QCab*, est mieux estimé que ces deux variables individuellement. En effet, les rôles de *LAI* et *Cab* sont individualisés dans les modèles de transfert radiatif, mais ce qui explique physiquement la réflectance du couvert est le contenu global en chlorophylle. De ce fait, on peut, par exemple, bien estimer *QCab* en sous-estimant *LAI* et en sur-estimant *Cab*. Il est important de prendre cette remarque en considération, notamment dans la façon dont on cherchera à appréhender la quantité d'azote absorbée par la plante, *QN*, à partir des variables biophysiques issues de la télédétection.

IV.2.2.2 Détermination de *LAI* et *QCab* à Chambry

Des mesures de réflectance du couvert ont été menées à Chambry de 2000 à 2003. Seules les mesures de 2000 et 2001 ont été exploitées pour notre étude, nous nous restreindrons donc à celles-ci. Elles ont été effectuées au cours de ces deux années par un capteur multi-spectral CASI par la société Astrium qui a fourni les variables *LAI* et *Cab* inversées selon leurs algorithmes proches de ceux développés à l'INRA et les données de réflectances.

Le Tableau IV.2-1 synthétise les dates de mesure pour les deux années concernées.

Parcelle 1	Parcelle 2
2000	2001
8 avril	2 avril
6 mai	21 avril
2 juin	8 mai
28 juin	31 mai

Tableau IV.2-1 : Dates d'acquisition des mesures de télédétection sur les parcelles 1 et 2 respectivement en 2000 et 2001.

Quelques remarques s'imposent au sujet de ces dates. Il était prévu d'acquérir des mesures de télédétection dès les mois de février et de mars, vu que, dans la perspective du

raisonnement de la fertilisation azotée qui doit être réalisé au mois de mars, des mesures à ces dates sont essentielles pour caractériser le couvert. Ceci n'a pas été rendu possible à cause des conditions climatiques, ces mesures nécessitant un ciel parfaitement dégagé. D'un côté, cela met en évidence une difficulté supplémentaire, pour le blé, de raisonner la fertilisation azotée à partir d'observations de télédétection de l'année en cours. Mais, d'un autre côté, conformément à ce qui a été dit sur la nécessité d'avoir des mesures pendant la phase de développement maximal de la culture (Chapitre III), les mesures effectuées sont assez bien situées temporellement.

Les estimations de *LAI* et *Cab* qui ont été utilisées pour notre étude ont été fournies par Astrium à une résolution de 2×2 m^2 et calibrées par des mesures au sol (Hollecker, 2002). La variable *QCab* a été obtenue en multipliant simplement *LAI* par *Cab*. Ces cartes ont ensuite été agrégées à l'échelle de pixels de 20×20 m^2 (P. Clastre, M. Pringle).

IV.2.3 Relations entre les données issues de la télédétection et les variables simulées par un modèle de culture

Grâce aux données de télédétection et à l'inversion des modèles de transfert radiatif, on estime donc le contenu en chlorophylle du couvert (*Cab* et *QCab*). Or, ce qui est simulé par le modèle est la quantité d'azote absorbé par la culture (*QN*). Il est donc nécessaire d'établir un lien entre ces deux quantités. Cet aspect a donné lieu à un important travail au cours de notre étude mais sera réduit ici à sa plus simple expression. Le menu détail en est donné dans Houlès *et al.* (soumis).

Nous développerons ici des éléments de bibliographie (i) qui justifient que l'on puisse prétendre estimer *QN*, variable simulée par le modèle de culture, à partir de *Cab* ou de *QCab*, estimée par télédétection et (ii) qui délimitent la confiance que l'on pourra avoir dans les relations empiriques qui vont être présentées, puis la façon dont ces relations ont pu être déterminées.

IV.2.3.1 Aspects de bibliographie

D'après Curran (1989) et Baret & Fourty (1997), l'azote n'absorbe de rayonnement que par ses liaisons avec des atomes de carbone, d'oxygène ou d'hydrogène. Il serait donc plus aisé de mesurer sa quantité via la mesure de la quantité de protéines, réservoir essentiel d'azote dans la cellule, que de le mesurer directement en tant qu'élément. Mais il est très mal aisé d'estimer la teneur en protéines par des mesures d'absorption ou de réflectance en raison

de la faiblesse de leur signal d'absorption (Fourty *et al.*, 1996). En revanche, la chlorophylle présente un fort pic d'absorption dans le domaine du visible et est très liée à la réflectance du couvert.

IV.2.3.1.1 Lien entre chlorophylle et azote dans la feuille

La chlorophylle est un élément essentiel de la biochimie des feuilles car elle conditionne la photosynthèse et, partant de là, l'ensemble du métabolisme cellulaire et notamment les protéines. On peut, très schématiquement, d'après ce constat, séparer l'azote contenu dans les feuilles en deux pools : celui qui est directement impliqué dans la photosynthèse et donc associé à la chlorophylle et celui qui est impliqué dans le reste du métabolisme. Hikosaka & Terashima (1996) indiquent un ordre de grandeur de chacun de ces pools : l'azote impliqué dans la photosynthèse (chlorophylle, antennes collectrices de la lumière, centres réactionnels, chaîne de transport des électrons, enzymes du cycle de Calvin et Benson dont la RUBISCO) représente environ 50% de l'azote total des feuilles. L'azote contenu dans les protéines non photosynthétiques représenterait environ 30% et les autres formes d'azote 20%. Lawlor *et al.* (1997) donnent des ordres de grandeur similaires : 55% pour l'azote impliqué dans la photosynthèse, 25% pour les protéines non photosynthétiques et 20% pour l'azote structural. Hikosaka & Terashima (1995) proposent un modèle, validé par eux-mêmes en 1996, donnant la répartition de l'azote entre les différents constituants de l'appareil photosynthétique. Celui-ci est basé sur des relations entre ces constituants s'apparentant à des relations stœchiométriques ; cette approche intéressante présente au demeurant deux limites : elle ne permet pas d'estimer le reste de l'azote contenu dans les feuilles ; de nombreux auteurs mettent en évidence l'extrême complexité des relations « stœchiométriques » liant les différents éléments de l'appareil photosynthétique. Les stress azoté et hydrique notamment affecteraient les relations entre divers constituants impliqués dans la photosynthèse (Macnab *et al.*, 1987 ; Sivanskar *et al.*, 1998 ; Lu & Zhang, 2000 ; Esquível *et al.* (2000) ; Karrou & Maranville, 1995 ; Shangguan *et al.*, 2000). Le Roux *et al.* (2001) rapportent une influence de la position de la feuille dans le couvert sur la répartition de l'azote disponible dans les différents pools photosynthétiques et d'autres auteurs indiquent également un gradient vertical de la concentration en azote dans le couvert qui pourrait résulter de différences dans la stœchiométrie photosynthétique (Shiraiwa & Sinclair, 1993 ; Lemaire & Gastal, 1997 ; Bindraban, 1999). Esquível *et al.* (2000), Karrou & Maranville (1995) et Shangguan *et al.* (2000) attestent également un effet du génotype sur la stœchiométrie, ce qui est une preuve de l'existence d'une régulation complexe, d'ordre

génétique, de celle-ci. Enfin, Macnab *et al.* (1987) observent un effet de la température sur la répartition de l'azote entre chlorophylle a et b et Osaki *et al.* (1993) montrent un effet de l'âge de la plante sur la répartition de l'azote entre les différents pools, différant selon le génotype.

Ce bref aperçu indique que la chlorophylle et l'azote sont bel et bien liés dans le métabolisme cellulaire : la quantité de chlorophylle conditionne l'efficacité de la photosynthèse et celle-ci conditionne le reste du métabolisme par sa fonction de producteur primaire. Mais il existe des régulations complexes entre les différents éléments de la chaîne. Il faut donc être conscient que les relations expérimentales pouvant être établies à l'échelle du couvert entre des teneurs en chlorophylle et en azote resteront hautement empiriques. En outre, la quantité d'azote contenue dans la chlorophylle est minime par rapport à l'azote total de la feuille. Par exemple, pour une concentration en chlorophylle a et b de 1.33 10^{-2} g g^{-1} et en azote de 4.17 10^{-2} g g^{-1} et pour un ratio de chlorophylle a sur b de 3, l'azote contenu dans la chlorophylle représente 2% de l'azote total (ces valeurs ont été extraites du jeu de données utilisé au cours de ce travail).

IV.2.3.1.2 Lien entre chlorophylle et azote à l'échelle du couvert

Une très grande majorité des travaux traitant de la relation entre teneur en chlorophylle (*Cab*) et teneur en azote (*N*) au niveau de la feuille ou du couvert remplace les premières par des mesures indirectes, de type *SPAD* (Minolta, 1990) ou Hydro-N-Tester (Hydro Agri, 2000), basées sur des mesures de transmittance. Ces mesures étant indirectes, il faut percevoir que la relation *SPAD* → *N* contient les deux sous-relations *SPAD* → *Cab* et *Cab* → *N*. Des Hydro-N-Tester ayant été utilisés au cours de notre travail pour réaliser les mesures au sol ayant servi à valider les estimations de *Cab* par télédétection, il convient d'en dire un mot. Des relations expérimentales *SPAD* → *Cab* ont également été étudiées à cette fin. Pour plus de simplicité, nous désignerons par *SPAD* toutes les mesures de ce type, indifféremment de la marque de l'appareil utilisé.

A l'origine, les mesures de type *SPAD* ont été conçues comme un moyen rapide et efficace pour raisonner la fertilisation azotée. Un grand nombre d'auteurs trouvent une bonne corrélation entre les mesures *SPAD* et les teneurs en azote de la plante (Piekelek & Fox, 1992 ; Feibo *et al.*, 1998 ; Vidal *et al.*, 1999 ; Reeves *et al.*, 1993 ; Matsunaka *et al.*, 1997), montrant ainsi la relation à l'échelle du couvert entre la teneur en chlorophylle et en azote. Mais, d'une étude à l'autre, ces relations sont hautement variables : elles dépendent de la variété (Užík & Žofajová, 2000), du stade de développement de la culture (Peng *et al.*, 1993 ;

Hoel, 1998 ; Shadchina *et al.*, 1998), de l'âge de la feuille (Castelli *et al.*, 1996 ; Sadras *et al.*, 2000), de la position de la mesure sur la feuille (Vidal & Bergareche, 1983 ; Hoel, 1998 ; Cartelat *et al.*, 2003) et de la position sur la plante de la feuille sur laquelle est effectuée la mesure, en raison des gradients verticaux de concentration en azote le long de la plante (Shirawai & Sinclair, 1993 ; Lemaire & Gastal, 1997 ; Bindraban, 1999). Les auteurs s'accordent en général sur le fait que les mesures doivent porter sur le deuxième tiers de la longueur des plus jeunes feuilles pleinement développées (Piekielek & Fox, 1992 ; Lemaire *et al.*, 1997 ; Feibo *et al.*, 1998 ; Vidal *et al.*, 1999). Plus marginalement, Hoel & Solhaug (1998) montrent que l'heure de la journée à laquelle la mesure *SPAD* est effectuée, par l'effet de la quantité de lumière sur la position des chloroplastes dans les cellules, affecte également les mesures. Shadchina *et al.* (1998) montrent cependant une certaine stabilité des relations d'une année sur l'autre pour un même stade de développement, alors que Dumont *et al.* (2001) suggèrent de travailler à l'échelle intégrée, c'est-à-dire de raisonner non pas en terme de concentration, mais en terme de quantité de chlorophylle et d'azote présente dans le couvert, ce qui permet de lever la difficulté soulevée par plusieurs auteurs qui ont montré (i) une meilleure corrélation entre les mesures *SPAD* et les teneurs surfaciques et non massiques (Peng *et al.*, 1993 ; Chapman & Barreto, 1997) et (ii) la difficulté d'estimer le terme de passage entre teneurs surfacique et massique (Roderick *et al.*, 1999) : la surface foliaire spécifique (Specific Leaf Area, SLA).

De l'analyse bibliographique exposée ci-dessus, il ressort qu'il est préférable de travailler sur la relation entre chlorophylle et azote (i) en mesurant réellement la teneur en chlorophylle et sans passer par les mesures *SPAD* qui constituent un intermédiaire introduisant des erreurs supplémentaires ; (ii) en prenant en compte l'effet du stade ; (iii) en travaillant à l'échelle du couvert, c'est-à-dire en déterminant des relations entre *QCab* et *QN*, contenus en chlorophylle et en azote du couvert, plutôt qu'entre teneurs. Ce dernier point est particulièrement important, étant donné que, comme on vient de le voir, *QCab* est mieux estimée que *Cab* par inversion de mesures de télédétection et qu'en outre le modèle STICS simule mieux des quantités absorbées (*QN*) que des concentrations.

IV.2.3.2 Détermination de relations expérimentales à Chambry

Afin de déterminer des relations expérimentales entre *QCab* et *QN* pour la variété *Shango*, des essais de doses d'azote ont été réalisés à Chambry en 2000 et 2001 (*cf. supra*). Des mesures de teneurs en azote des plantes étaient effectuées selon la méthode Dumas

(Ebeling, 1968), par combustion totale (analyseur élémentaire EA-Isochrone). Des mesures de teneur en chlorophylle foliaire ont été effectuées d'après la méthode proposée par Moran (1982), basée sur l'extraction de la chlorophylle des tissus par du *N,N*-Dimethylformamide, de préférence à la méthode proposée par Arnon (1949), beaucoup plus courante (e.g. Vidal & Bergareche, 1983 ; Shadchina *et al.*, 1998 ; Le Roux *et al.*, 2001), utilisant de l'acétone et nécessitant un broyage des feuilles coûteux en temps. Après extraction, l'absorbance du solvant est mesurée à différentes longueurs d'ondes avec un spectrophotomètre (Beckman-DU 64). Des relations entre la teneur en chlorophylle du solvant et son absorbance à certaines longueurs d'onde sont proposées notamment par Lichtenthaler & Wellburn (1983). Nous avons également établi des relations pour le *N,N*-Dimethylformamide grâce à une gamme étalon. Les valeurs de *QCab* pour les différents traitements d'azote ont été établies en multipliant ces mesures de *Cab* foliaire par les mesures de *LAI* et celles de *QN* en multipliant les teneurs en azote des plantes par leur biomasse.

Le Tableau IV.2-2 présente les dates de mesures réalisées les deux années avec, en vis-à-vis, les dates d'acquisition d'images de télédétection les plus proches. Les prélèvements n'ont pas toujours été effectués le jour des acquisitions, notamment en 2001 : dans ce cas, les dates les plus proches sont utilisées, en vérifiant que sur la période les relations sont confondues.

2000		**2001**	
Télédétection	**Relations *QCab-QN***	**Télédétection**	**Relations *QCab-QN***
08/04	06/04	02/04	28/03 et 11/04
06/05	06/05	21/04	25/04
02/06	02/06	08/05	02 et 17/05
28/06	28/06	31/05	31/05

Tableau IV.2-2 : Correspondance entre les dates des acquisitions des images de télédétection et des prélèvements de végétaux ayant permis d'établir les relations entre *QCab* et *QN*.

La Figure IV.2-1 présente les relations entre *QCab* et *QN* utilisées pour chaque date d'acquisition de données de télédétection de l'année 2000 et la Figure IV.2-2 pour 2001. Ces relations, pour une date donnée, sont de bonne qualité, avec une dispersion assez limitée, ce qui montre l'intérêt et la pertinence de l'approche au niveau du couvert (coefficients de régression proches ou supérieurs à 0.9), sauf la dernière date en 2000 qui a été établie avec moins de points.

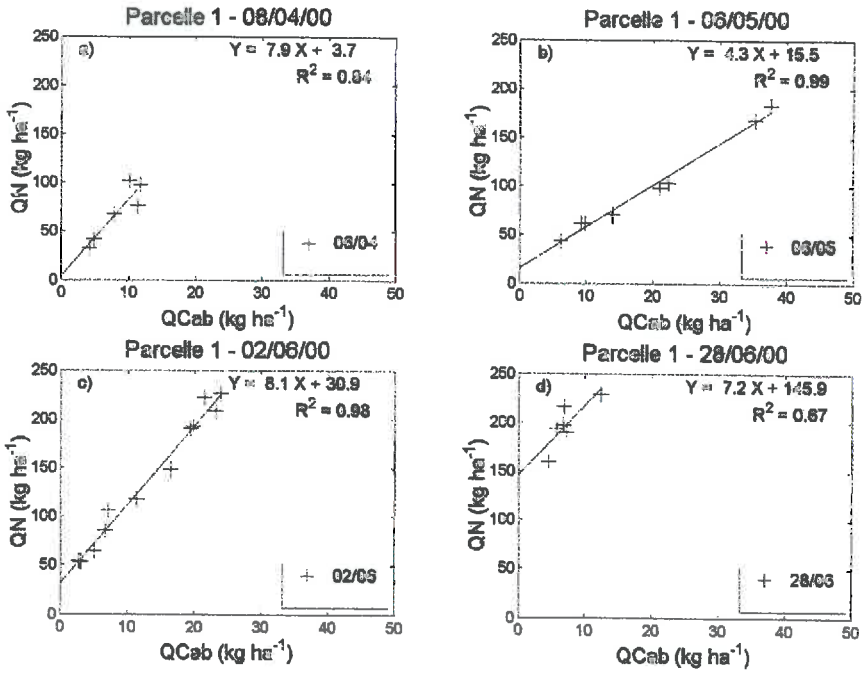


Figure IV.2-1 : Relations expérimentales établies à différentes dates au cours de l'année 2000 et servant à convertir les valeurs de quantité de chlorophylle dans le couvert (*QCab*) obtenue par télédétection en quantité d'azote absorbé par la culture (*QN*) utilisée par le modèle. Les droites de régression sont indiquées ainsi que les coefficients de régression.

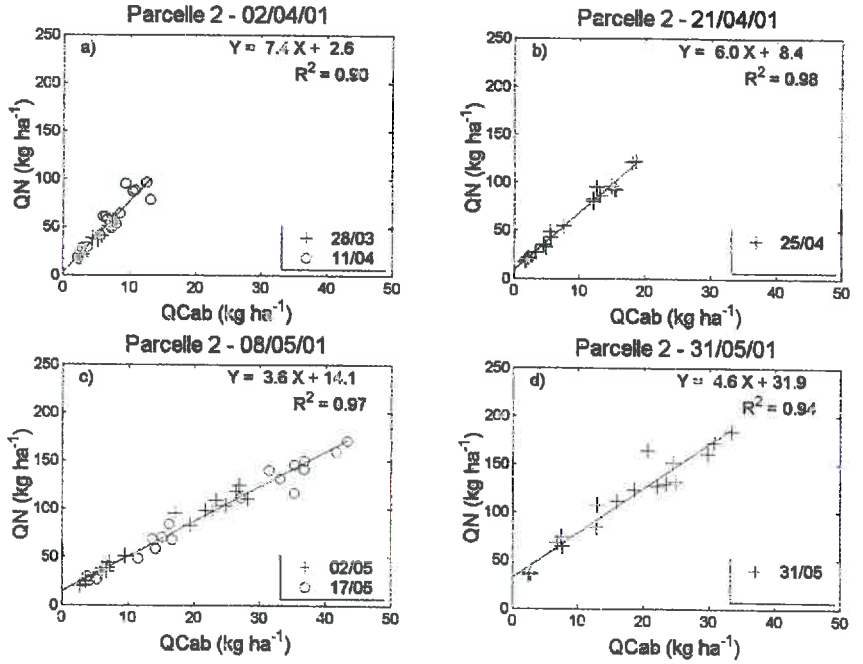


Figure IV.2-2 : Relations expérimentales établies à différentes dates au cours de l'année 2001 et servant à convertir les valeurs de quantité de chlorophylle dans le couvert (*QCab*) obtenue par télédétection en quantité d'azote absorbé par la culture (*QN*) utilisée par le modèle. Pour le 02/04 et le 08/05, des mesures expérimentales n'ayant pas été effectuées, les deux dates de mesures les plus proches ont été utilisées. Les droites de régression sont indiquées ainsi que les coefficients de régression.

Les coefficients des ces relations peuvent être paramétrés en fonction du stade de développement, exprimé par exemple en degrés jours, de façon à pouvoir être utilisés pour d'autres années et à n'importe quelle date.

Des cartes caractérisant la variabilité spatiale de *QN* ont ainsi pu être établies à partir de celles de *QCab* : elles vont être présentées avec les cartes de *LAI* dans le paragraphe suivant.

IV.2.4 Visualisation des cartes établies à Chambry

Cette partie permet de visualiser et de comparer les différentes cartes dont l'établissement vient d'être décrit. Les relations entre les valeurs de *LAI*/*QN* mesurées en cours de culture et de rendement seront également commentées.

IV.2.4.1 Variables estimées via la télédétection : *LAI* et *QN*

La Figure IV.2-4 présente les cartes de *LAI* et *QN* estimées au cours de quatre acquisitions en 2000 sur la parcelle 1. La Figure IV.2-5 fait de même avec la parcelle 2 en 2001.

Pour la parcelle 1, on constate d'une part une très grande hétérogénéité à faible échelle des variables estimées : différentes zones sont identifiables mais restent de taille modeste. La plus grande d'entre elles apparaît en marron sur toutes les cartes vers le sud-est de la parcelle et est caractérisée par les plus fortes valeurs de *LAI* et *QN* observées. On observe du reste une relative stabilité dans le temps des structures spatiales.

Pour la parcelle 2, de nombreux pixels situés en bordure sont caractérisés par des valeurs aberrantes, particulièrement faibles. Un point noir met en évidence ceux dont la valeur d'au moins une des variables observées est inférieure à la moyenne de la parcelle diminuée de 1.25 fois l'écart type. D'une date à l'autre, les points isolés selon ce critère sont très constants. Nous verrons dans la suite les conséquences de la mauvaise estimation de la valeur de ces pixels. On peut attribuer ces mauvaises mesures à des effets de bordure, qui peuvent autant avoir des causes agronomiques (tassement du sol par exemple) que physiques (pollution des pixels de bord de champ par les pixels voisins, très différents). La structure spatiale de la parcelle semble par ailleurs plus changeante au cours du temps qu'elle ne l'était pour la parcelle 1.

La Figure IV.2-3 représente l'évolution des valeurs moyennes de *LAI* et de *QN* avec des indications sur leur dispersion. Ces cinétiques sont assez cohérentes avec celles évaluées par des mesures destructives sur les essais de doses d'azote ayant reçu le même traitement que la parcelle. On retrouve le fait que certaines valeurs en 2001 sont particulièrement faibles. En outre, cette figure fait bien apparaître qu'en 2000 le cycle du *LAI* est décrit jusqu'à un stade de sénescence avancé, alors qu'en 2001 les mesures s'arrêtent au plateau, ce qui ne pourra pas permettre, dans l'optique de l'assimilation, de caractériser les propriétés affectant la fin du cycle.

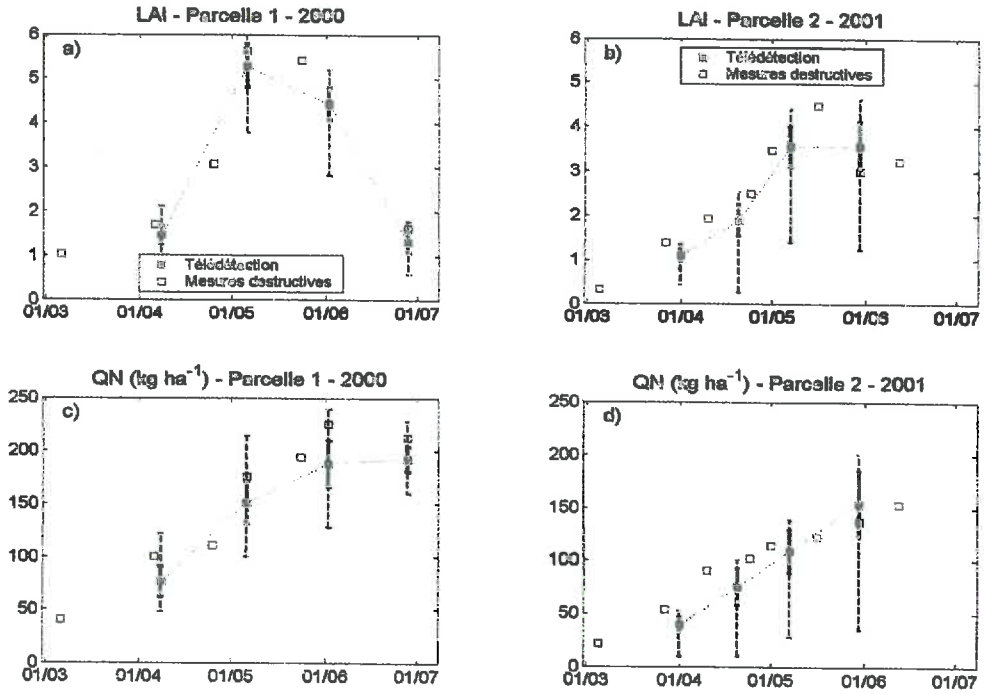


Figure IV.2-3 : Cinétiques des valeurs moyennes sur la parcelle de *LAI* et *QN* sur la parcelle 1 en 2000 et la parcelle 2 en 2001 (carrés noirs). a) *LAI* estimé en 2000 ; b) *LAI* estimé en 2001 ; c) *QN* estimée en 2000 ; d) *QN* estimée en 2001. Les barres d'erreur en traits pleins indiquent les écarts types et les barres d'erreur en traits pointillés les valeurs minimales et maximales. A titre de comparaison, les valeurs de *LAI* et de *QN* estimées par mesure destructive sur les essais de doses d'azote ayant reçu la même fertilisation azotée que la parcelle entière sont indiquées par un carré blanc.

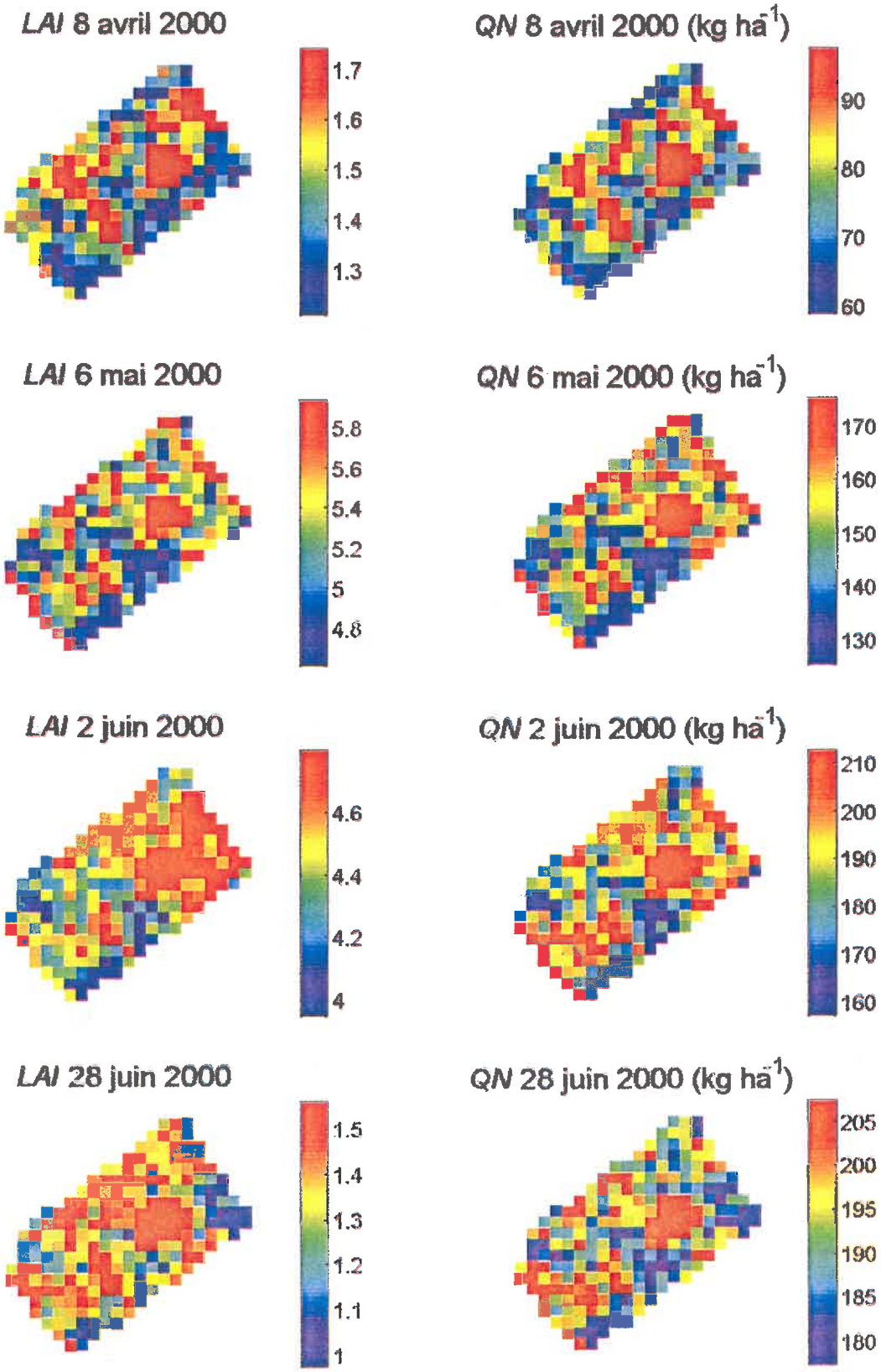


Figure IV.2-4 : Cartes des variables biophysiques *LAI* et *QN* dérivées de mesures de télédétection à différentes dates sur les parcelles de Chambry en 2000. Résolution : 20×20 m^2.

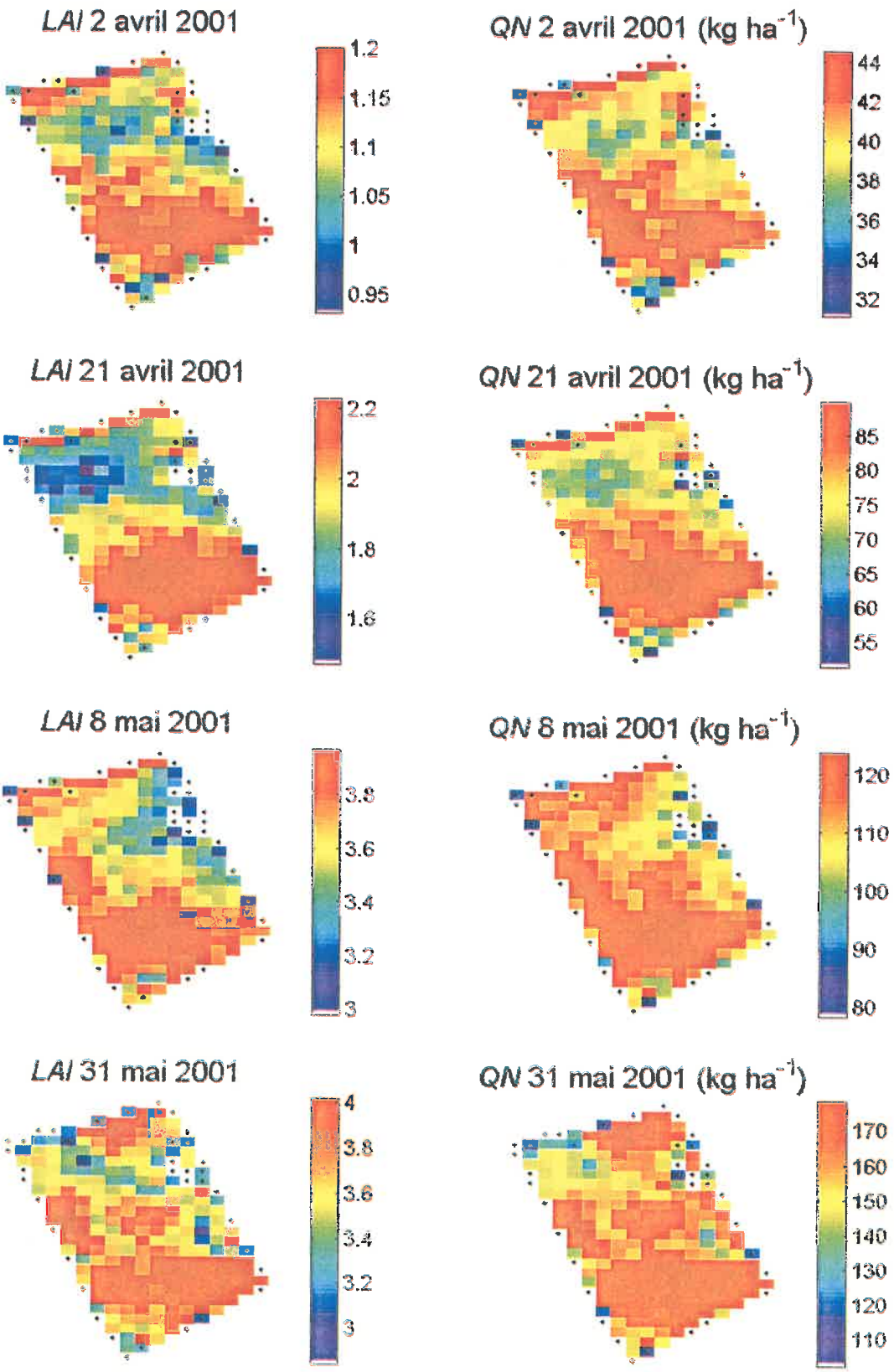


Figure IV.2-5 : Cartes des variables biophysiques *LAI* et *QN* dérivées de mesures de télédétection à différentes dates sur les parcelles de Chambry en 2001. Résolution : 20×20 m^2. Les points noirs marquent les pixels pour lesquels au moins une des huit variables représentées est inférieure à la moyenne de la parcelle diminuée de 1.25 fois l'écart type.

IV.2.4.2 Variables estimées à la récolte : rendement et teneur en protéines

La Figure IV.2-6 permet de visualiser les cartes de rendement et de teneur en protéines. On constate une relative faiblesse de la variabilité intra-parcellaire : la gamme de variation du rendement se situe autour de 1 t ha^{-1} ; elle est plus forte en 2002 (1.5 t ha^{-1}), année caractérisée par un traitement azoté en carré latin. La variation de la teneur en protéines est très faible en 2000 et 2001, années où la fertilisation azotée était uniforme : moins de 0.3% de protéines. Pourtant, les quantités d'azote absorbées étaient beaucoup plus variables : 30 kg ha^{-1} fin juin 2000 et presque 80 kg ha^{-1} fin mai 2001. Les années où le traitement n'était pas uniforme par contre, la teneur en protéines à la récolte est beaucoup moins homogène : on atteint une amplitude de 3% en 2002, ce qui est principalement dû à la variabilité du quatrième apport (le 6 mai) et de la quantité totale apportée (120 kg ha^{-1} d'amplitude). En 2003, où le traitement n'était pas non plus homogène, on observe aussi une assez importante variabilité : plus de 1.5% de protéines. Ceci illustre le fait que sur ces parcelles pour le moins, la variabilité de la teneur en protéines est beaucoup plus liée à la dose d'azote qu'à la variabilité du sol. Visuellement, l'effet du traitement azoté en carré latin en 2002 est bien perceptible et ce surtout sur les protéines, ce qui confirme ce qui vient d'être dit. Il est donc difficile de comparer la variabilité spatiale concernant la parcelle 1 d'une année sur l'autre à cause du traitement azoté différencié de 2002. Pour la parcelle 2, des structures communes apparaissent entre 2001 et 2003 : notamment, sur les cartes de rendement, une zone de forte production au sud-ouest et une zone de faible production sur la bordure est. Les cartes de teneur en protéines en revanche n'ont pas du tout le même aspect : en 2001, où la variabilité est très faible, elle ressemble fortement à la carte des rendements. En 2003, où la variabilité est plus forte en raison du traitement azoté, la carte présente des similitudes avec la carte des sols présentée dans le Chapitre III.

Il est difficile de trouver des structures communes entre les cartes de *LAI* et de *QN* et ces variables caractéristiques de la récolte. La partie suivante s'attache à étudier la corrélation entre les valeurs de rendement et de *LAI* et *QN* mesurées en cours de culture.

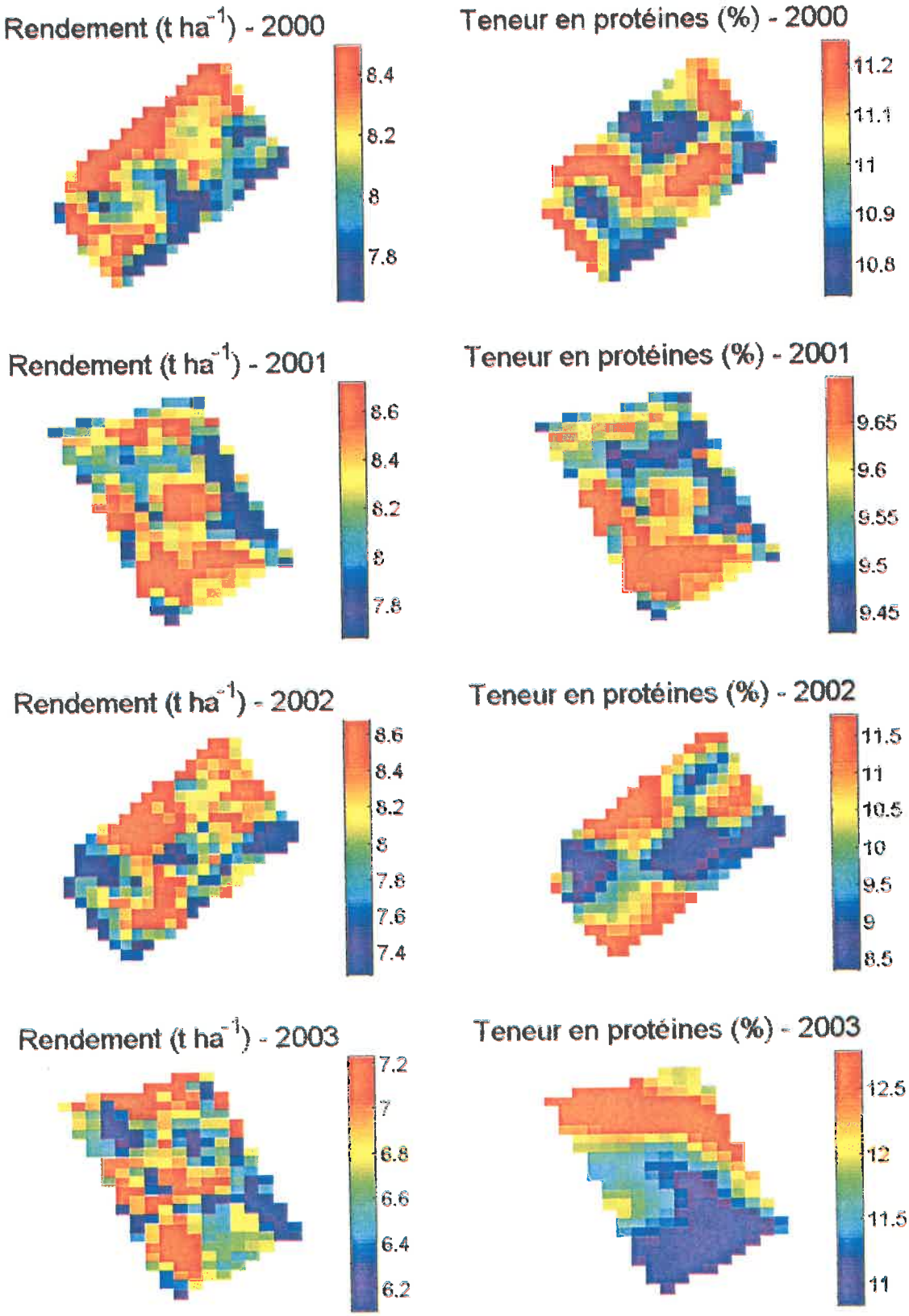


Figure IV.2-6 : Cartes de rendement et de teneur en protéines mesurées à la récolte en 2000, 2001, 2002 et 2003. Résolution : 20x20 m^2.

IV.2.4.3 Corrélations entre *LAI*, *QN* et rendement

La Figure IV.2-7 représente les relations entre le rendement mesuré et les estimations de *LAI* et de *QN* aux quatre dates de mesures. Le Tableau IV.2-3 donne les valeurs des coefficients de corrélation de ces variables. Ces variables sont assez faiblement liées, notamment en 2001 où le meilleur coefficient de corrélation n'est que de 0.41. Pour les deux parcelles, les plus forts coefficients sont ceux qui sont obtenus avec les dates les plus tardives de mesure, ce qui est cohérent avec le fait que, d'après le modèle, les dates les plus informatives sont celles de la fin du cycle (Chapitre III). Ceci fait apparaître qu'il est difficile de prévoir le rendement directement à partir de mesures de *LAI* et de *QN* : pour ce faire, les connaissances synthétisées dans un modèle de culture constituent un apport essentiel.

Il convient enfin de remarquer qu'une partie de la mauvaise relation entre ces variables est liée au fait qu'elles ont été évaluées par des moyens très différents.

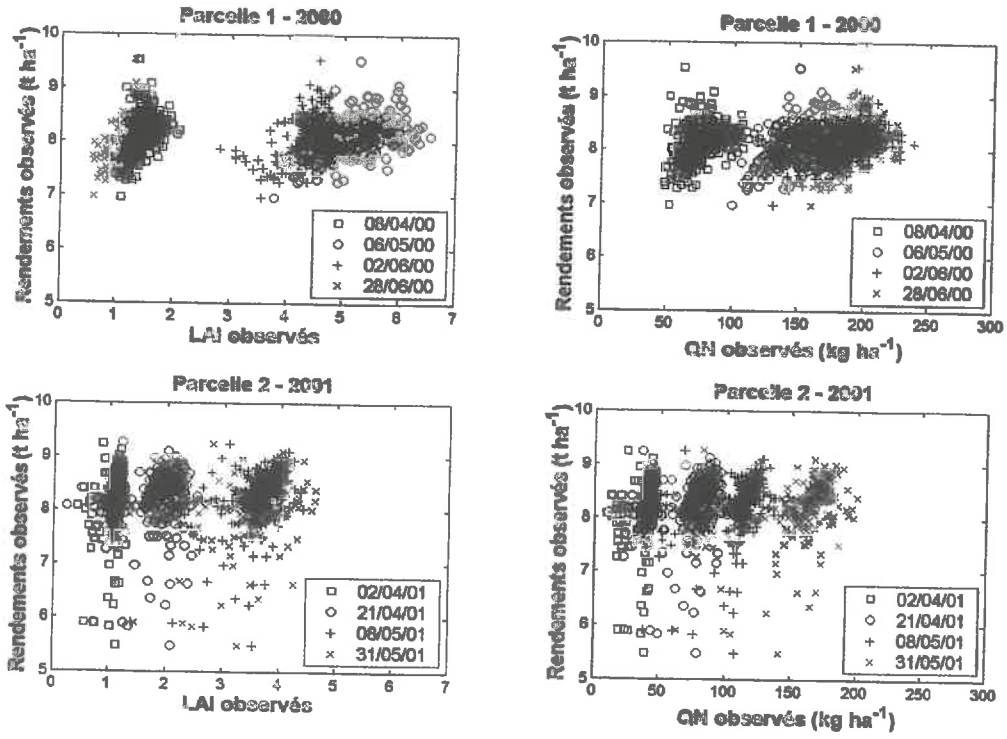


Figure IV.2-7 : Relations entre le rendement mesuré à la récolte et les valeurs de *LAI* et *QN* observées en cours de culture pour les parcelles 1 et 2 en 2000 et 2001 respectivement. Les valeurs des coefficients de corrélation sont indiquées dans le Tableau IV.2-3.

	Parcelle 1 -2000	Parcelle 2 - 2001
LAI date 1	0.42	0.28
LAI date 2	0.47	0.24
LAI date 3	0.39	0.39
LAI date 4	0.56	0.41
QN date 1	0.34	0.28
QN date 2	0.43	0.32
QN date 3	0.46	0.37
QN date 4	0.53	0.36

Tableau IV.2-3 : Valeurs des coefficients de corrélation entre les valeurs de *LAI* et *QN* observées en cours de culture à différentes dates et le rendement mesuré à la récolte pour les parcelles 1 et 2 en 2000 et 2001 respectivement.

IV.2.5 Conclusion

L'ensemble des cartes présentées va servir à évaluer l'aptitude des deux méthodes de spatialisation du modèle à décrire la variabilité de la culture à l'échelle intra-parcellaire. Les variables *LAI* et *QN* serviront essentiellement à faire l'assimilation et à porter un diagnostic sur la caractérisation de la variabilité en cours de culture grâce à l'approche par cartographie. Le rendement et, dans une moindre mesure, la teneur en protéines serviront à comparer les deux approches sur leur aptitude à correctement représenter les variables qui sont impliquées dans le calcul de la dose d'azote optimale.

IV.3 Comparaison de l'approche par cartographie et de l'approche par assimilation

Cette partie correspond à la mise en œuvre sur les jeux de données réelles des deux approches développées au cours du Chapitre III. Leurs performances en terme de qualité de simulation des variables impliquées dans le calcul du critère de détermination de la dose optimale d'azote seront comparées sur les quatre années de mesure, de 2000 à 2003. Le premier paragraphe étudie les différences entre les échelles d'acquisition des différentes variables impliquées dans cette comparaison et en tire les conséquences par rapport au choix d'une résolution commune avant de décrire la mise en œuvre des deux méthodes de spatialisation. Les cinq paragraphes qui viennent ensuite traitent : (i) de la simulation de *LAI* et *QN* en 2000 et 2001 ; (ii) de la caractérisation des propriétés du sol et des plantes par les deux approches ; (iii) de la simulation du rendement et de la teneur en protéines en 2000 et 2001 ; (iv) de la simulation du rendement et de la teneur en protéines en 2002 et 2003 ; (v) d'une méthode d'évaluation de la qualité de la caractérisation de la variabilité spatiale par chacune des deux approches. Enfin, une discussion permettra de dégager les points essentiels de cette comparaison.

IV.3.1 Mise en œuvre de la comparaison

IV.3.1.1 Echelle d'acquisition des données et de comparaison

Les échelles auxquelles ont été acquises les différentes données nécessaires à la mise en œuvre des deux approches de spatialisation et à leur validation sont fort différentes. Le Tableau IV.3-1 les synthétise en précisant l'origine des mesures. Vu cette hétérogénéité, il n'était pas trivial de fixer une échelle commune pour la comparaison, ce choix pouvant pour partie conditionner les résultats. Les nœuds de grille auraient pu être choisis car ils correspondent à l'échelle d'établissement de l'approche par inventaire cartographique des sols qui sert de référence. Mais cette échelle pose problème pour la carte de rendement, qui est acquise à une échelle moins ponctuelle : il aurait été imprécis d'attribuer à un nœud de grille la valeur correspondante de la carte de rendement. Un problème analogue se posait pour les données issues de mesures de télédétection. Bien que ce puisse être sujet à discussion et sans présager d'études ultérieures, il a été résolu de choisir l'échelle des pixels de 20×20 m^2 comme support commun de toutes les données. Cela permet un certain compromis entre le

temps de simulation nécessaire, beaucoup plus faible que si nous avions travaillé à l'échelle des pixels de télédétection de 2×2 m^2, et l'impossibilité d'étudier la variabilité spatiale de façon continue si nous avions travaillé à l'échelle des nœuds de grille.

Approche de spatialisation	Types de données	Origine des acquisitions	Echelle d'acquisition	Spatialisation (20×20 m^2)
	LAI et *QN*	Inversion télédétection + relations *QCab – QN*	2×2 m^2	Agrégation
	Rendement	Carte de rendement	6×5 m2	Agrégation
	Teneur en protéines	Grille	Ponctuelle	Krigeage
Cartographie	Caractéristiques permanentes	Grille et fosses pédologiques	Ponctuelle	UCS
	Eau et azote au semis	Grille	Ponctuelle	Krigeage
Assimilation	Variables d'entrée estimées	Assimilation de *LAI* et de *QN*	20×20 m^2	Déjà fait
	Variables d'entrée non estimées	Moyenne des mesures grille	Parcelle	Aucune

Tableau IV.3-1 : Origine, échelle d'acquisition et passage à l'échelle de comparaison de 20×20 m^2 par type de données et par type d'approche de spatialisation de chacune des deux approches de spatialisation.

Pour effectuer ce passage, deux méthodes ont été utilisées selon le type de données : l'agrégation qui consiste simplement à moyenner les pixels regroupés pour les données acquises de façon continue dans l'espace à une résolution plus précise que la résolution commune ; le krigeage pour les données acquises ponctuellement sur les points de la grille. Cette façon de faire ne permet pas de réaliser facilement des comparaisons en tout point de la parcelle, c'est pourquoi nous n'avons guère utilisé par la suite les teneurs en protéines pour valider les simulations. La spatialisation des caractéristiques permanentes de sol dans l'approche par cartographie a été réalisée d'une façon particulière comme expliqué dans le Chapitre III en utilisant la notion d'UCS qui implique à la fois une expertise et des observations d'images aériennes de la parcelle. Une fois la couverture établie, elle a été scindée en pixels de 20×20 m^2.

IV.3.1.2 Mise en œuvre des deux méthodes de spatialisation

Approche par cartographie : une fois les UCS et UTS définies, il suffit d'appliquer à chaque UTS les règles de pédotransfert (Chapitre III). Ceci permet de renseigner pour chaque pixel de 20×20 m^2 les caractéristiques permanentes des sols. Il faut ensuite faire de même avec les cartes d'eau et d'azote au semis obtenues par krigeage. Pour effectuer ces simulations, tous les horizons ont été ramenés à une épaisseur de 30 cm.

Approche par assimilation : la méthode GLUE a été employée pour réaliser l'assimilation sur les années 2000 et 2001 des variables *LAI* et *QN* estimées à la résolution des pixels de 20×20 m^2. N_G=200 000 simulations de départ ont été effectuées. Les variables estimées par assimilation de *LAI* et de *QN* sont les mêmes que celles qui ont été dégagées au cours du Chapitre III : *CALC*, *DA(2)*, *HCCF(1)*, *HCCF(2)*, *HCCF(3)*, *NO3initf(1)*, *NORG*, *OBSTARAC*, *q0*, *durvieF* et $vlai_{max}$. Pour les commentaires relatifs à ce choix, nous invitons le lecteur à se reporter à ce chapitre. *ARGI* a cependant été supprimée car elle faisait double emploi avec *CALC* (*cf.* Chapitre II). L'information *a priori* sur ces variables correspondait aux valeurs minimales et maximales mesurées sur chaque parcelle. Celle des paramètres plante *durvieF* et $vlai_{max}$ étaient respectivement [140 ; 220] et [1.5 ; 2.5]. Toutes les autres variables de STICS ont reçu la même valeur en tout point. Les variables sol non estimées ont été fixées à la moyenne des mesures effectuées sur la grille. Les estimations des erreurs des variables observées étaient proportionnelles à leurs valeurs selon un coefficient de $k_{LAI} = 0.15$ pour *LAI* et $k_{QN} = 0.20$ pour *QN*. Ces coefficients ont été estimés d'après des données de validation des inversions (Moulin, comm. pers.). Les quatre dates d'observation présentées dans le Tableau IV.2-1 ont été utilisées. Les distributions *a posteriori* déterminées ont permis de réaliser des simulations des variables d'intérêt après assimilation. Les années 2002 et 2003 serviront à valider les assimilations faites respectivement en 2000 et 2001.

IV.3.2 Comparaison des simulations du *LAI* et de *QN* en 2000 et 2001

La Figure IV.3-1 représente les valeurs de *LAI* et de *QN* simulées par les approches par cartographie et par assimilation. Les valeurs de *LAI* et *QN* observées sur ces graphiques correspondent à celles qui ont été utilisées pour l'assimilation : il ne s'agit donc pas ici de réaliser une validation de cette approche, mais de vérifier qu'elle permet bien de diminuer l'erreur de simulations sur ces variables, ce qui est un pré-requis.

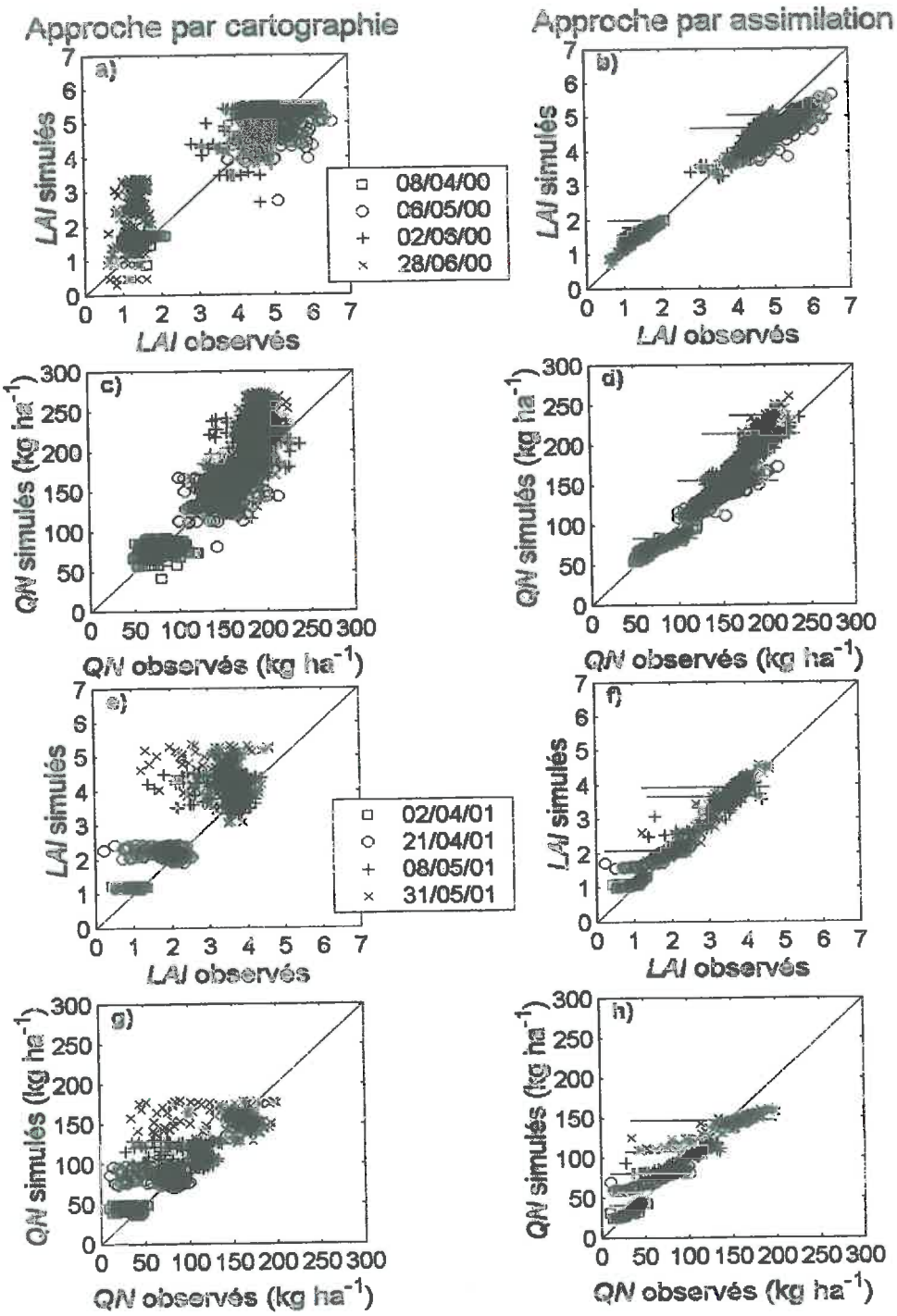


Figure IV.3-1 : Simulations en 2000 (a à d) par l'approche par cartographie de *LAI* (a) et de *QN* (c) et par l'approche par assimilation avec la méthode GLUE de *LAI* (b) et de *QN* (d). Simulations en 2001 (e à h) par l'approche par cartographie de *LAI* (e) et de *QN* (g) et par l'approche par assimilation avec la méthode GLUE de *LAI* (f) et de *QN* (h) aux quatre dates de mesure. Dans les graphiques concernant l'approche par assimilation, les simulations données par l'information *a priori* sont indiquées par un trait horizontal. Il y a 280 points par date en 2000 et 268 en 2001, soit le nombre de pixels des parcelles 1 et 2 respectivement.

Le Tableau IV.3-2 donne les valeurs de RMSE totale, relative et décomposées en partie systématique et non-systématique pour *LAI* et le Tableau IV.3-3 pour *QN*.

LAI	2000			2001		
	Cartographie	IAP	Assimilation	Cartographie	IAP	Assimilation
RMSE	0.83	0.53	0.37	0.73	0.46	0.20
RRMSE	27%	17%	12%	29%	18%	8%
RMSEs	0.53	0.42	0.32	0.41	0.22	0.08
RMSEu	0.63	0.34	0.20	0.60	0.40	0.19

Tableau IV.3-2 : RMSE, RRMSE, RMSEs et RMSEu entre les valeurs de *LAI* observées et simulées dans l'approche par cartographie, par assimilation ou à partir de l'IAP utilisée par cette dernière pour les années 2000 et 2001. Unités (sauf RRMSE) : m^2 de feuilles par m^2 de sol.

QN	2000			2001		
	Cartographie	IAP	Assimilation	Cartographie	IAP	Assimilation
RMSE	33	31	15	24	21	14
RRMSE	22%	21%	10%	26%	22%	15%
RMSEs	17	21	5	12	12	10
RMSEu	28	23	14	21	17	11

Tableau IV.3-3 : RMSE, RRMSE, RMSEs et RMSEu entre les valeurs de *QN* observées et simulées dans l'approche par cartographie, par assimilation ou à partir de l'IAP utilisée par cette dernière pour les années 2000 et 2001. Unités (sauf RRMSE) : kg ha^{-1}.

On constate que l'approche par cartographie restitue généralement assez bien en moyenne les valeurs de *LAI* et de *QN* observées, ce qui est indiqué par la position centrée autour de la première bissectrice des nuages aux différentes dates. Ceci n'est cependant pas le cas des dates les plus tardives, surtout en 2000, ce qui témoigne d'une sénescence simulée pas assez précoce. En revanche, la variabilité spatiale n'est pas bien restituée : pour les dates les plus tardives, les nuages de points ont une forme très arrondie, ce qui se traduit par des valeurs de RMSEu qui sont plus fortes que les valeurs de RMSEs. Pour les dates les plus précoces, (08/04/00 ; 02/04/01 ; 21/04/01), les valeurs simulées sont très regroupées : cela montre que le modèle ne permet pas de restituer une variabilité spatiale en début de cycle de culture sur la seule base d'une variabilité des propriétés du sol et justifie l'emploi dans le processus d'assimilation de paramètres plante comme *durvieF* et $vlai_{max}$.

Les simulations basées sur l'assimilation permettent en revanche de reproduire cette variabilité spatiale précoce. En 2001 on voit cependant qu'on atteint une valeur limite minimale en deçà de laquelle le modèle semble ne pas pouvoir aller. Mais ces points observés particulièrement faibles correspondent précisément aux points de bordure de parcelle. En

2000, on observe sur *LAI* un biais assez fort le 06/05, ce qui explique la valeur de RMSEs plus forte en 2000 qu'en 2001. Mais, globalement, l'assimilation a permis de bien améliorer la simulation de *LAI* et *QN* par rapport à l'IAP, ce qui était attendu.

IV.3.3 Comparaison des variables d'entrée

Trois grands types de variables d'entrée peuvent être comparées : celles qui jouent sur le stress hydrique, celles qui jouent sur le stress azoté et celles qui sont des paramètres plante, $vlai_{max}$ et *durvieF*. Le premier type peut être comparé de façon synthétique en calculant les réserves utiles. Les autres seront comparées en terme de distribution sur la parcelle.

IV.3.3.1 Comparaison de l'estimation de la réserve utile

En utilisant les variables d'entrée du sol estimées par les deux approches, une variable plus synthétique, la réserve utile, *RU*, a été calculée de la façon suivante :

$$RU = \int_0^{OBSTARAC} [HCCF(z) - HMINF(z)] \cdot DA(z) \cdot dz \qquad \text{Équation IV.3-1}$$

On a donc considéré que la réserve utile était la quantité d'eau disponible pour la plante entre la surface du sol et l'obstacle à l'enracinement. Dans cette approche, le caractère dynamique de la réserve utile lié à l'accroissement du front racinaire n'est pas pris en compte : il s'agit d'une potentialité du sol. En ce qui concerne l'approche par assimilation, la RU que l'on calcule est une RU « équivalente », qui traduit plus le fonctionnement hydrique du sol vis à vis de la plante qu'elle ne représente une RU au sens strict du terme. Par exemple, elle peut englober l'effet des remontées capillaires qui ne sont pas simulées par le modèle.

La Figure IV.3-2 présente cette comparaison sous forme de carte et de graphique. Le premier constat qui peut être fait est que l'on retrouve très peu de structures communes dans la distribution spatiale de la réserve utile entre les deux façons de la décrire. Dans la parcelle 1, des taches se retrouvent cependant d'une carte à l'autre alors que dans la parcelle 2, le gradient nord-sud caractéristique de l'approche par cartographie des sols ne se retrouve pas du tout dans l'approche utilisant l'assimilation. Les cartes de RU obtenues par assimilation présentent naturellement de grandes similitudes avec celles de *LAI* et de *QN*.

Les graphiques font apparaître l'absence de corrélation entre les deux descriptions issues des deux approches. Sur la parcelle 1, l'approche par cartographie distingue trois grandes zones distinctes (correspondant aux trois nuages) alors que l'approche par assimilation caractérise la variabilité de façon beaucoup plus continue et indépendante de ces

zones. Cette dernière donne des valeurs de RU globalement plus faibles, ce qui va dans le sens d'une augmentation des stress simulés.

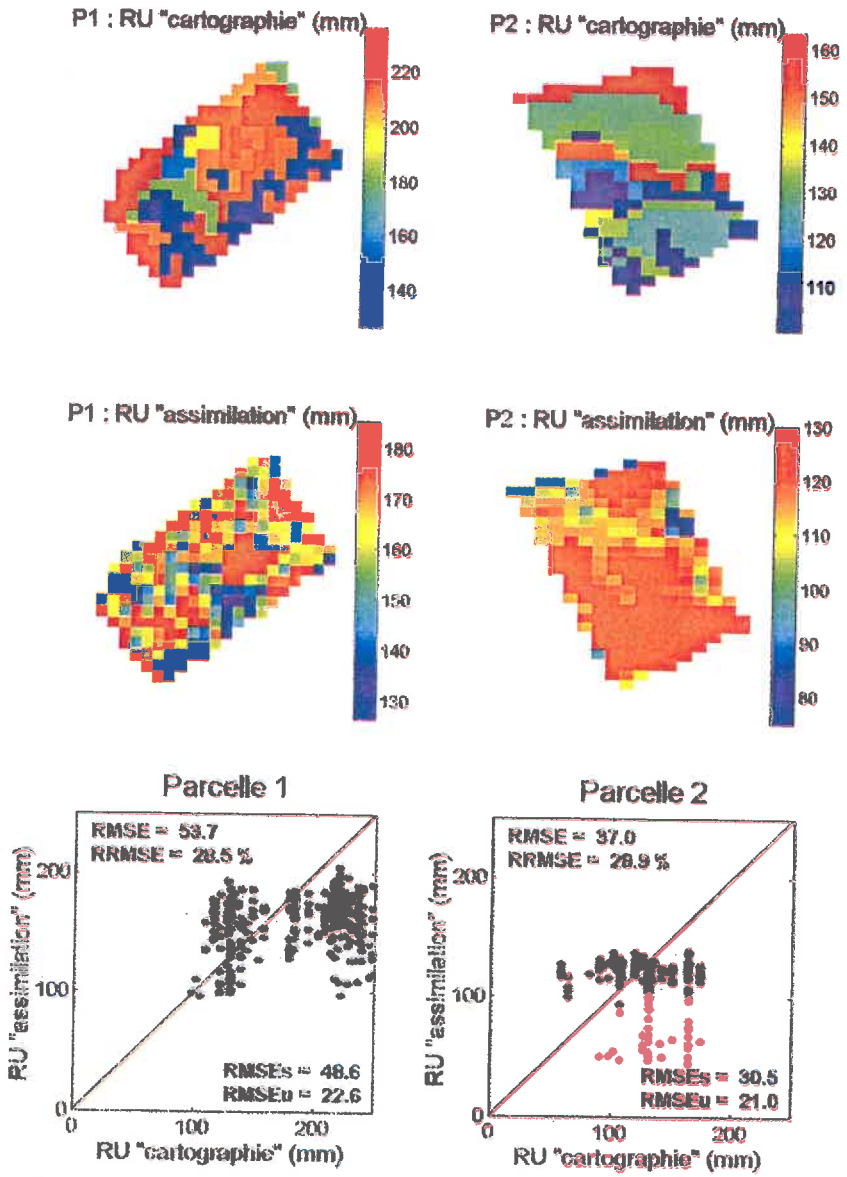


Figure IV.3-2 : Comparaison de l'estimation des réserves utiles en utilisant l'approche par cartographie (RU « cartographie ») et l'approche par assimilation (RU « assimilation »), pour les parcelles 1 et 2. Les points en rouge dans le graphique concernant la parcelle 2 correspondent aux points de bordure caractérisés par de faibles valeurs de *LAI* et *QN* observées.

Sur la parcelle 2, si l'approche par cartographie distingue 5 à 6 zones bien distinctes, l'approche par assimilation, mis à part les pixels correspondant aux faibles valeurs de *LAI* et *QN* observées, donne une très faible variation centrée autour de 130 mm environ. Cette valeur correspond au milieu de l'intervalle de variation donné par l'approche par cartographie et qui correspond à l'IAP : ceci signifie qu'en 2001, pour la majorité des points, la moyenne de la distribution *a posteriori* est proche de celle de la distribution *a priori*. Les conditions

climatiques ayant régné au cours de cette année particulièrement humide mis à part un bref épisode sec peuvent être à l'origine de la difficulté de réaliser cette caractérisation. Cependant, il est évidemment difficile de dire qu'une approche est meilleure que l'autre en l'absence de mesures réelles exhaustives des propriétés du sol.

IV.3.3.2 Comparaison de l'estimation des variables d'entrée sol jouant sur les stress azotés

Nous présentons ici une comparaison des distributions à la parcelle de variables d'entrée qui jouent sur l'azote disponible pour la plante, *NORG*, *CALC*, *OBSTARAC* et *NO3initf(1)*. La Figure IV.3-3 donne cette comparaison pour la parcelle 1 et l'année 1999-2000. La Figure IV.3-4 fait de même pour la parcelle 2 et l'année 2000-2001. Les valeurs extrêmes des axes des abscisses représentent les bornes de l'information *a priori*.

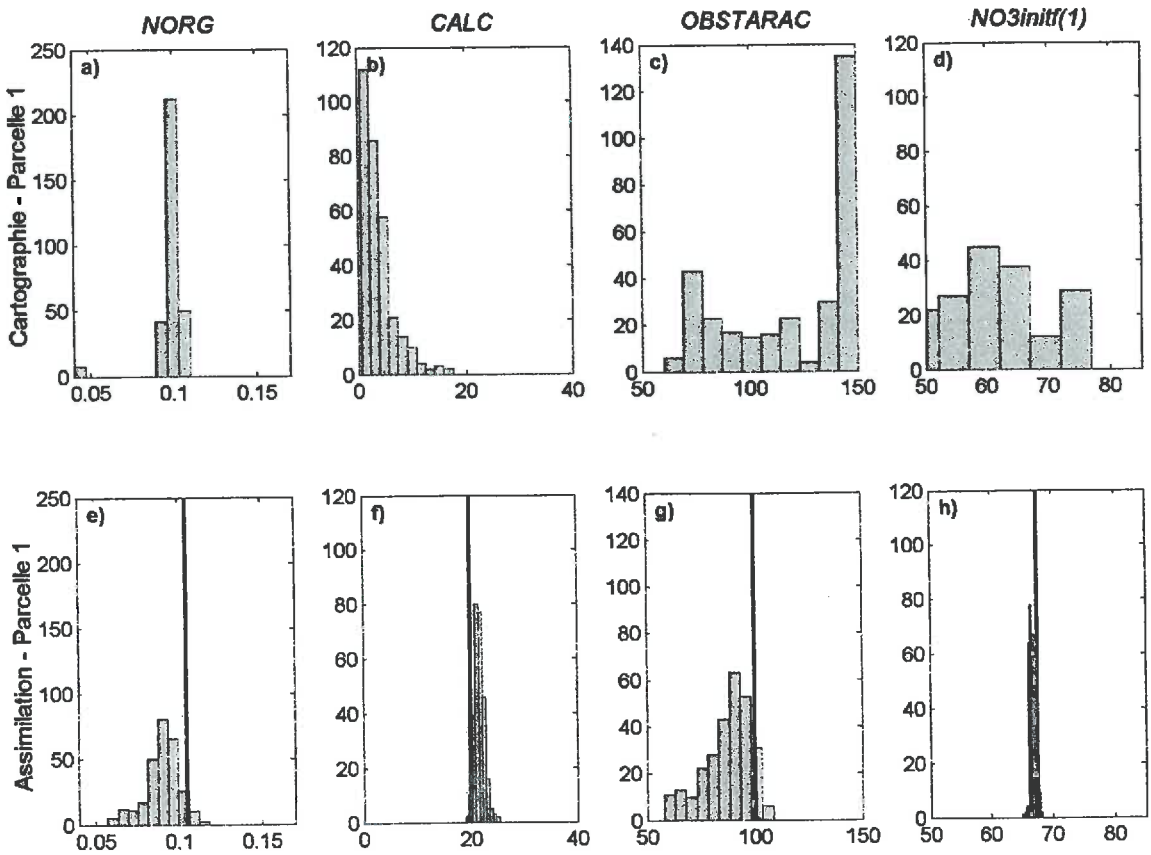


Figure IV.3-3 : Comparaison de l'estimation des variables d'entrée du modèle caractéristiques du sol *NORG* (% pondéral) (a et e), *CALC* (% pondéral) (b et f), *OBSTARAC* (cm) (c et g) et *NO3initf(1)* (kg ha^{-1}) (d et h) par l'approche par cartographie (a à d) et par assimilation (e à h) pour la parcelle 1, année climatique 1999-2000. Les valeurs extrêmes des axes des abscisses correspondent à l'intervalle définissant l'information *a priori* pour chacune des variables et la barre verticale en marque le centre.

Pour la parcelle 1, les valeurs de *NORG* et *OBSTARAC* obtenues par l'approche par assimilation sont plus faibles que celles obtenues par l'approche par cartographie et inversement pour *CALC*, ce qui se traduit par une augmentation des stress (*cf.* Tableau II.3-3) et une diminution du rendement simulé (*cf.* Figure III.3-10 et 11). *NO3initf(1)* est en revanche en moyenne un peu plus fort, ce qui se traduirait par une augmentation du rendement, mais avec un hiver pluvieux, cette variable n'a pas de poids. Par ailleurs, on constate que les valeurs de *CALC* et *NO3initf(1)* sont très regroupées autour du centre de l'intervalle défini par l'IAP. Ceci illustre ce qui avait été commenté au Chapitre III, à savoir que la méthode GLUE réalise implicitement une sorte d'analyse de sensibilité au sens adopté pour la méthode MCIP : elle n'estime que les variables permettant d'améliorer les simulations, les autres étant fixées à la valeur moyenne de l'intervalle. Les écarts types sur ces variables sont d'ailleurs très forts : leur distribution *a posteriori* correspond à leur distribution *a priori*. Les fortes valeurs de *CALC* proviennent donc essentiellement de la définition de l'IAP.

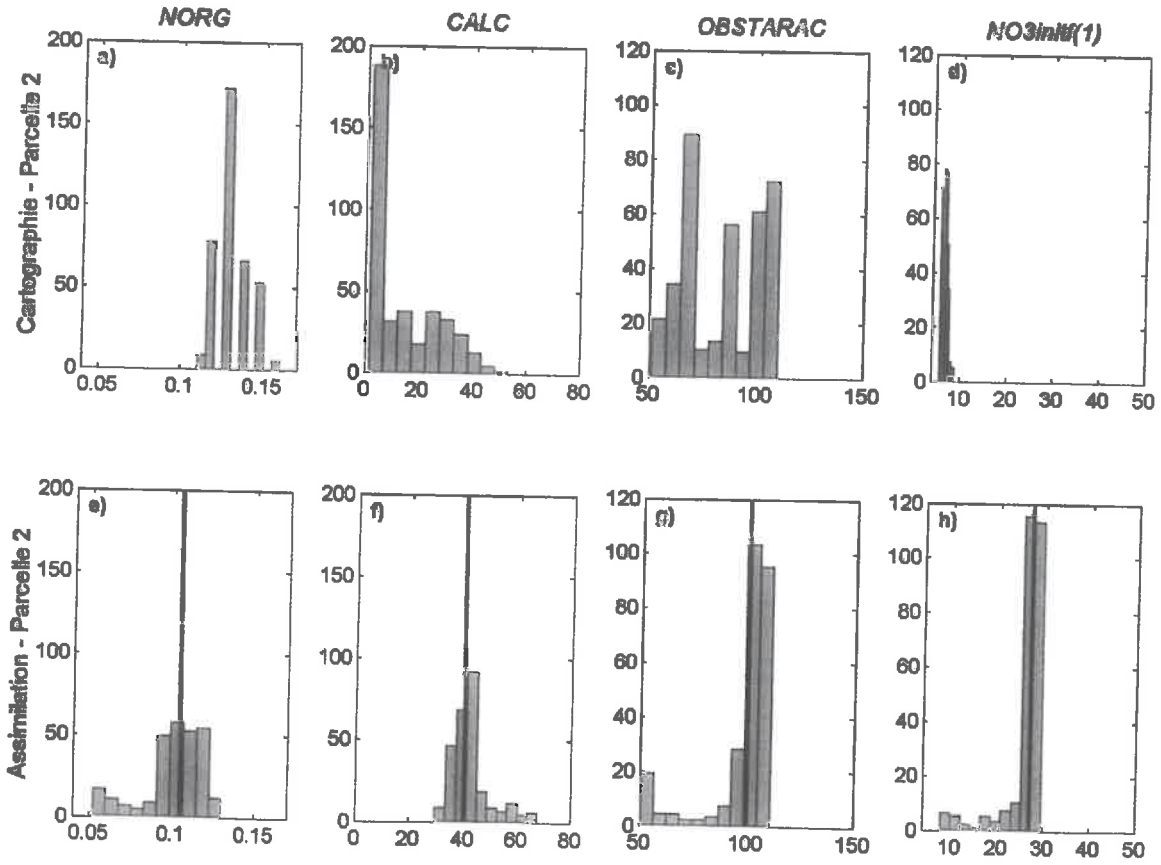


Figure IV.3-4 : Comparaison de l'estimation des variables d'entrée du modèle caractéristiques du sol *NORG* (% pondéral) (a et e), *CALC* (% pondéral) (b et f), *OBSTARAC* (cm) (c et g) et *NO3initf(1)* (kg ha^{-1}) (d et h) par l'approche par cartographie (a à d) et par assimilation en utilisant la méthode GLUE (e à h) pour la parcelle 2, année climatique 2000-2001. Les valeurs extrêmes des axes des abscisses correspondent à l'intervalle définissant l'information *a priori* pour chacune des variables et la barre verticale en marque le centre.

Pour la parcelle 2, les valeurs obtenues par assimilation de *NORG* et *CALC* par rapport à celles obtenues par cartographie auront tendance à faire diminuer le rendement, alors que les valeurs de *OBSTARAC* et *NO3initf(1)* auront l'effet inverse. En outre, les distributions à la parcelle sont, pour ces quatre variables, très centrées autour du centre de l'IAP, mis à part une dispersion qui correspond aux points de bordure de parcelle caractérisés par de faibles valeurs de *LAI* et *QN* observées. Ceci montre que la qualité de la simulation de *LAI* et *QN* était peu sensible à ces variables et qu'elles n'ont pas été estimées, comme c'était le cas des variables déterminant la RU. Seule la variable *NORG* se démarque de cette analyse, présentant une dispersion plus forte.

IV.3.3.3 Comparaison de l'estimation des variables d'entrées correspondant à des paramètres plante

L'utilisation des paramètres plante *durvieF* et $vlai_{max}$ dans l'assimilation était justifiée par (i) la nécessité de faire entrer ce genre de paramètres dans le processus d'assimilation pour permettre de mieux simuler *LAI* comme montré par une étude préliminaire et (ii) l'hypothèse qu'ils pouvaient traduire des interactions sol-plante non prises en compte par le modèle (carence phospho-potassique, compaction par exemple).

Pour les deux années, ces deux paramètres plante ont été réestimés avec une variabilité géographique marquée (Figure IV.3-5). Leur évolution par rapport aux valeurs utilisées dans l'approche par cartographie tend à faire diminuer les valeurs de *LAI* simulées. La diminution de *durvieF* a également un fort impact sur la simulation du rendement et de la teneur en protéines par le raccourcissement de la période d'accumulation de matière sèche (*cf.* Chapitre II). Ceci pourra expliquer directement les faibles rendements simulés en 2001 et par suite en 2003. En 2001, ce sont donc essentiellement ces paramètres qui ont permis de spatialiser les simulations, avec *NORG*, car on a vu que la plupart des autres variables d'entrées estimées par assimilation n'avaient pas changé de distribution sauf pour les pixels de faibles valeurs observées de *LAI* et *QN* situés en bordure de parcelle.

IV.3.4 Comparaison des variables de sortie en 2000 et 2001

Dans cette partie nous nous attacherons surtout à décrire la qualité de la simulation du rendement. Etant donnée la façon dont les valeurs de teneur en protéines ont été acquises, la comparaison est, on l'a vue, plus difficile. Nous dirons également un mot du bilan d'azote et de la marge brute dont le calcul se fait à partir du rendement et de la teneur en protéines.

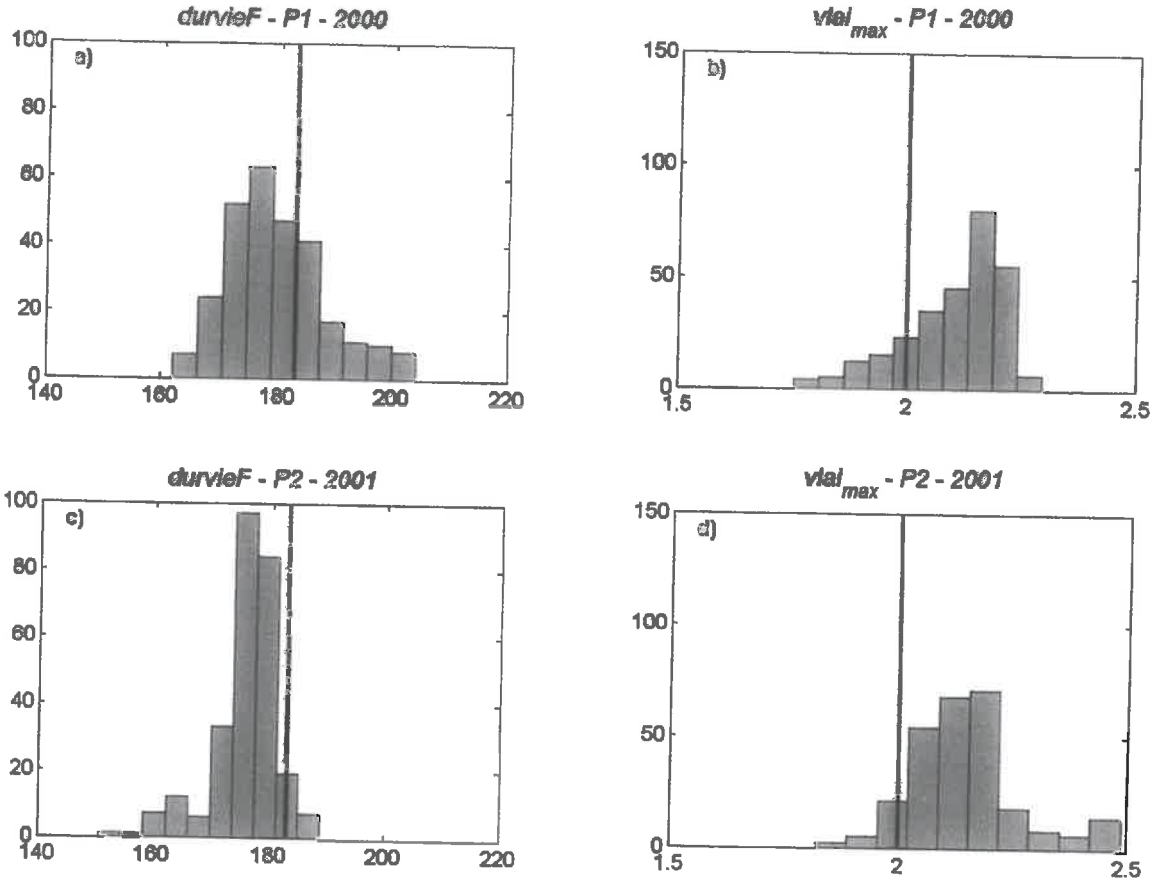


Figure IV.3-5 : Comparaison de l'estimation des variables d'entrée du modèle correspondant à des paramètres plante, *durvieF* (a et c) et $vlai_{max}$ (b et d) par assimilation en utilisant la méthode GLUE aux valeurs standard utilisées dans l'approche cartographique (trait vertical) pour la parcelle 1, année climatique 1999-2000 (a et b) et pour la parcelle 2, année climatique 2000-2001 (c et d). Les valeurs extrêmes des axes des abscisses correspondent à l'intervalle donné par l'information *a priori* pour chacune des variables.

IV.3.4.1 Comparaison des valeurs observées et simulées

La Figure IV.3-6 présente les résultats de simulation de rendement obtenus en 2000 et 2001 avec les approches par cartographie et par assimilation. Le Tableau IV.3-4 donne les valeurs de RMSE correspondantes et leur décomposition.

G	2000			2001				
	Cartographie	IAP	Assimilation	Cartographie		IAP	Assimilation	
RMSE	2.27	1.80	1.24	1.70	1.74*	2.33	2.67	2.19*
RRMSE	28%	22%	15%	21%	21%*	28%	32%	27%*
RMSEs	2.31	1.14	1.14	1.11	1.22*	2.33	2.33	2.15*
RMSEu	0.77	0	0.49	1.30	1.24*	0	1.02	0.43*

Tableau IV.3-4 : RMSE, RRMSE, RMSEs et RMSEu entre les valeurs de rendement observées et simulées dans l'approche par cartographie, par assimilation ou à partir de l'IAP utilisée par cette dernière pour les années 2000 et 2001. * : valeurs obtenues en 2003 après retrait des pixels de bordure. Unités (sauf RRMSE) : t ha^{-1}.

Les valeurs de rendement sont mal simulées par l'approche cartographique, de façon cohérente avec ce que nous venons de dire sur *LAI* : le nuage de points n'est pas aligné sur la première bissectrice et la variabilité simulée est très supérieure à la variabilité observée. Elles sont caractérisées par une forte dispersion (RMSEu=0.77 t ha^{-1} en 2000 et 1.30 en 2001). En 2000, où *LAI* et *QN* sont plutôt sur-estimés surtout en fin de cycle, le rendement l'est aussi : ceci démontre un biais dans la caractérisation spatiale des propriétés du sol par cette approche ou une mauvaise adéquation entre la façon de décrire un sol dans une approche pédologique et dans le modèle. En 2001, *LAI* était également plutôt sur-estimé alors que dans la plupart des zones, le rendement est sous-estimé. Ceci est directement attribuable aux valeurs très faibles (de 50 à 90/100 cm) affectées à *OBSTARAC* dans les zones sud de la parcelle notamment. Ces valeurs d'*OBSTARAC* peuvent ne pas affecter la simulation de *LAI* en cours de culture, voire compenser d'autres variables le favorisant trop. Mais elles peuvent en revanche diminuer les valeurs simulées des variables en fin de cycle comme le rendement, en exagérant le stress hydrique au cours du remplissage des grains. Ceci illustre le poids prépondérant de cette variable sur le rendement dans le modèle.

Les valeurs de rendement obtenues par l'approche basée sur l'assimilation font office de validation de cette méthode. En 2000, les simulations obtenues sur la base de l'information *a priori* sont en moyenne très surestimées et l'assimilation permet de retrouver des valeurs d'un ordre de grandeur plus conforme à celui des rendements observés. Le nuage de points est, comme l'étaient ceux du *LAI*, assez bien étiré sur la première bissectrice, ce qui montre que la spatialisation est relativement bien réalisée par cette approche. Pour les plus forts rendements cependant, on a une tendance à une surestimation qui va croissant. Les valeurs de RMSE témoignent de l'amélioration des simulations par rapport à l'IAP et d'une meilleure description que l'approche par cartographie (Tableau IV.3-4).

L'année 2001 met en évidence plusieurs problèmes. La spatialisation obtenue sur *LAI* se retrouve dans le gros du nuage qui est assez bien étiré selon une parallèle à la première bissectrice. Quelques points s'égaillent en dehors de ce nuage et correspondent tous aux valeurs particulièrement faibles de *LAI* observées correspondant aux pixels situés en bordures de carte. Les mauvaises performances sur ces points peuvent être entièrement attribuées à la mauvaise précision des mesures de *LAI* en bordure de parcelle avec toutes les conséquences que l'on a vu sur l'estimation des variables d'entrée. Les points caractérisés par un faible rendement observé et un rendement simulé moyen peuvent correspondre à des erreurs de

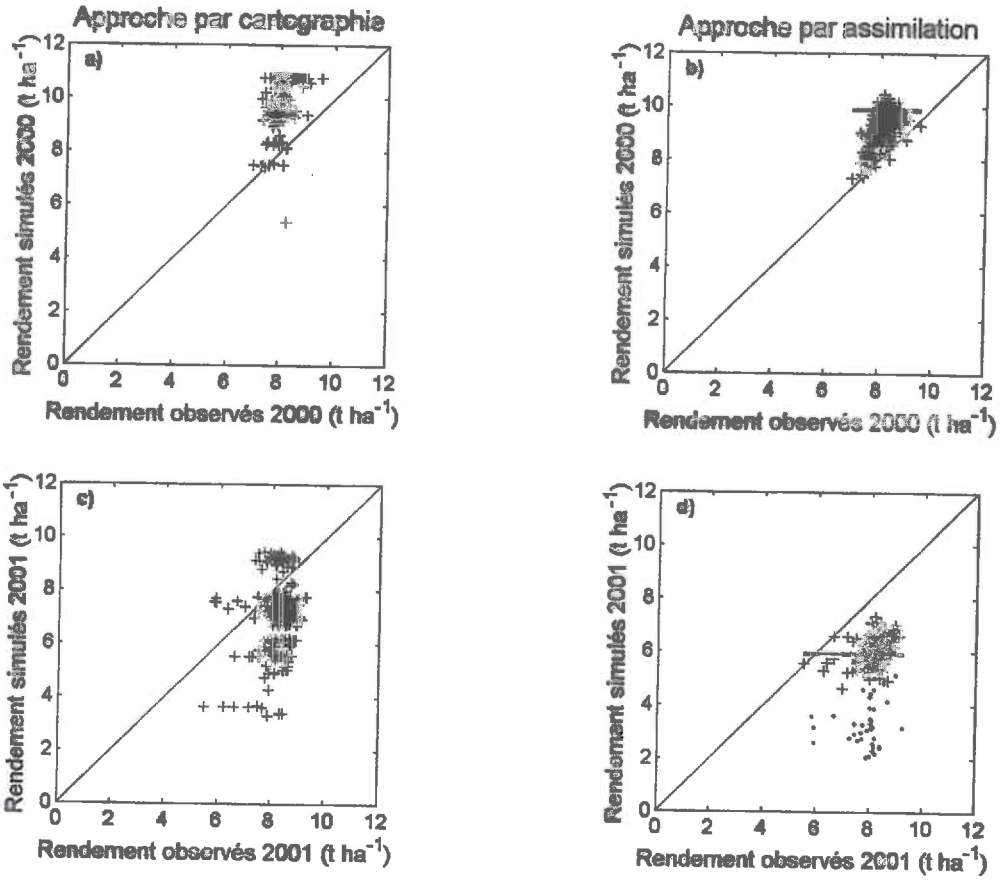


Figure IV.3-6 : Comparaison des simulations de rendement par l'approche par cartographie (a et c) et par l'approche par assimilation (b et d) pour les années climatiques 1999-2000 (a et b) et 2000-2001 (c et d). Dans les graphiques concernant l'approche par assimilation, les simulations données par l'information *a priori* sont indiquées par un trait horizontal.

mesure du rendement. L'approche par assimilation présente pour cette année également une plus faible dispersion que l'approche par cartographie, surtout après avoir supprimé les points de bordure de parcelle (Tableau IV.3-4). Pour le reste des points, un fort biais entache la qualité des simulations. Il peut être essentiellement dû aux assez faibles valeurs adoptées par *durvieF*. On remarque que les simulations après assimilation restent groupées dans les environs des simulations avant assimilation matérialisées par un trait vertical. On pourrait donc attribuer cette sous-estimation systématique au fait que l'information *a priori* ne permet pas de réaliser des simulations englobant les bonnes simulations mais ce n'est pas le cas. On est dans une situation où le modèle simule fort convenablement *LAI* et *QN* mais pas le rendement. Nous reviendrons sur cet aspect dans la discussion concernant cette partie.

La simulation de la teneur en protéines surestime également largement la variabilité observée. On constate de plus une forte corrélation négative entre cette variable et le rendement qui n'est pas constatée dans les données observées. Cette corrélation est liée à un

effet de concentration dans les grains de l'azote absorbé pour les faibles rendement, ce qui nous renvoie à des problèmes d'absorption de l'azote en fin de cycle notamment. La qualité de simulation du bilan d'azote et de la marge brute ressemble très fortement à celle du rendement.

IV.3.4.2 Etude des résidus sur le rendement, *LAI* et *QN*

La Figure IV.3-7 présente les relations entre les différences des valeurs de rendement observées (G) et simulées ($\hat{G}$) et les différences des valeurs observées et simulées de *LAI* et *QN* pour les deux approches. Les valeurs de résidus de *LAI* et *QN* sont normalisées entre 0 et 1. Deux aspects majeurs se dégagent : les résidus de G et de *QN* sont assez fortement corrélés dans les deux approches ($\rho \approx 0.8$) et ce, pour les deux années. Par contre, on a une perte de la corrélation entre les résidus de G et de *LAI* dans le cas de l'approche par assimilation, alors qu'elle existe encore dans le cas de l'approche par cartographie. Ceci illustre que dans le cas de l'approche par cartographie, il existe une certaine cohérence entre les erreurs commises sur G et sur *LAI*, alors que cette cohérence est perdue dans l'approche par assimilation : on constate une tendance à surestimer G associée à une tendance à sous-estimer *LAI* (faible corrélation négative). Ceci provient d'une part du fait que dans cette dernière approche les résidus calculés sur *LAI* sont très faibles, ce qui ne peut que faire baisser la corrélation. D'autre part, ceci pourrait être l'indicateur d'une perte de la cohérence entre les valeurs de variables d'entrée estimées par assimilation. En effet, dans ce processus, les corrélations éventuelles entre variables d'entrée n'ont pas été prises en compte, alors que les mesures de l'approche par cartographie respectent de fait ces liaisons entre les caractéristiques de sol. Enfin, on peut dire que les causes des erreurs commises sur le rendement ne sont pas liées dans l'approche par assimilation à des erreurs faites sur *LAI*. Dans le cas de *QN*, la corrélation entre résidus reste : cela indique que l'assimilation sur *QN* a été moins efficace que sur *LAI*.

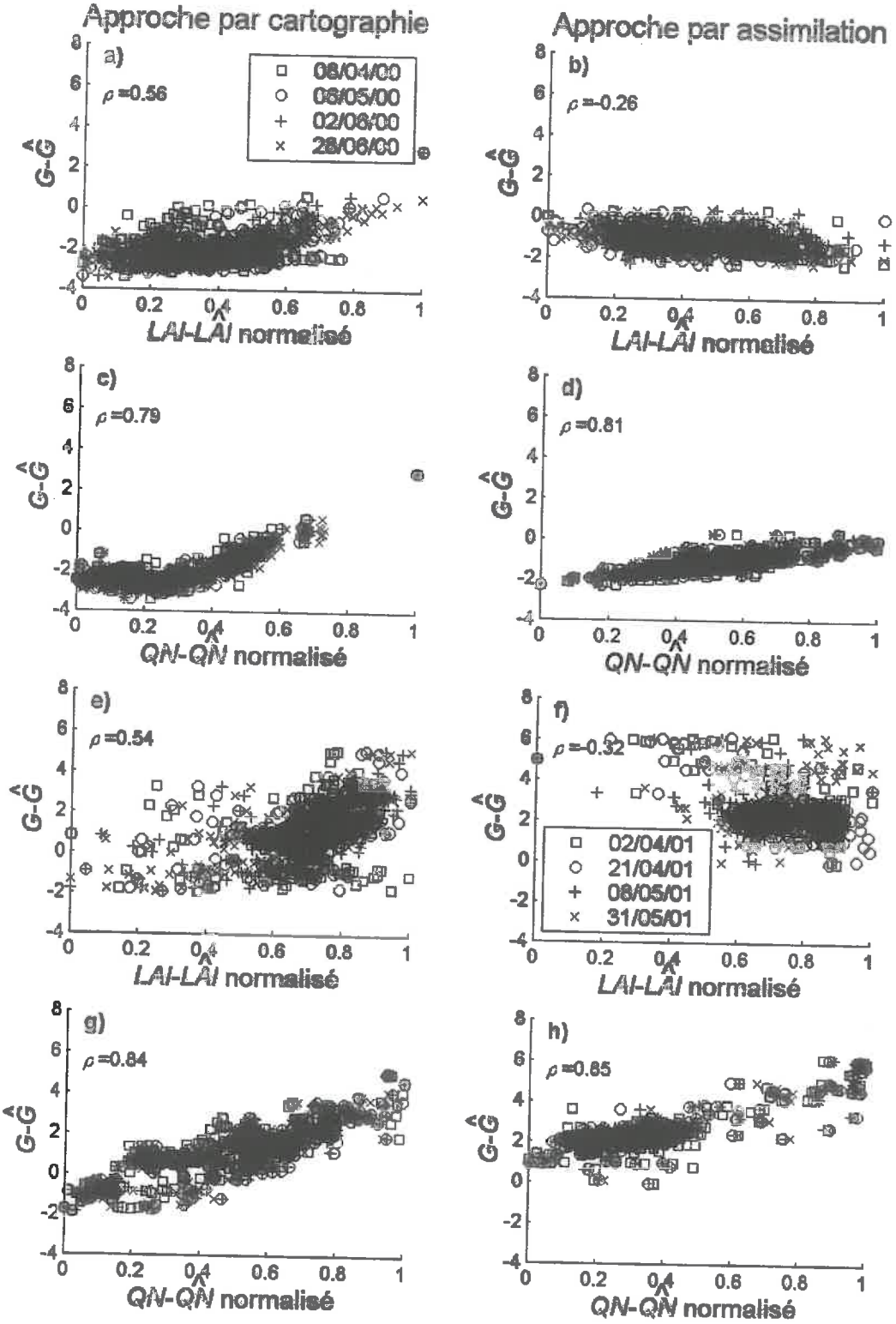


Figure IV.3-7 : Pour les deux approches de spatialisation et pour l'année 2000 (a à d) et 2001 (e à h), relation entre les différences des valeurs de rendement observées (G) et simulées ($\hat{G}$) et les différences des valeurs observées et simulées de LAI et QN normalisées entre 0 et 1 (les valeurs simulées sont notées $\hat{LAI}$ et $\hat{QN}$). Les valeurs des coefficients de corrélation (ρ) calculées toutes dates confondues sont indiquées.

IV.3.5 Comparaison des variables de sortie en 2002 et 2003

Nous rappelons à ce stade qu'au cours de l'année 2001-2002, les deuxième et troisième apports ont été modulés sur la parcelle 1 selon un carré latin comprenant neuf zones et trois doses différentes. Les doses totales apportées étaient 140, 200 et 260 kg ha^{-1} respectivement. L'année 2002-2003 a vu quant à elle un dispositif de fertilisation azotée modulée à l'est et homogène à l'ouest de la parcelle. Enfin, l'engrais azoté utilisé au cours de ces deux campagnes était sous forme solide, contrairement aux deux années précédentes où il était liquide, et caractérisé par une moindre efficience (*cf.* Chapitre I).

La Figure IV.3-8 présente les résultats de simulation de rendement obtenus en 2002 et 2003 avec les approches par cartographie et par assimilation. Le **Erreur ! Source du renvoi introuvable.** donne les valeurs de RMSE correspondantes et leur décomposition.

G	2002			2003				
	Cartographie	IAP	Assimilation	Cartographie		IAP	Assimilation	
RMSE	1.02	0.91	2.80	3.13	3.22	3.38	3.71	3.37
RRMSE	13%	11%	35%	47%	48%	51%	56%	51%
RMSEs	0.65	0.91	2.75	2.87	2.98	3.38	3.63	3.35
RMSEu	0.78	0	0.52	1.25	1.21	0	0.78	0.40

Tableau IV.3-5 : RMSE, RRMSE, RMSEs et RMSEu entre les valeurs de rendement observées et simulées dans l'approche par cartographie, par assimilation ou à partir de l'IAP utilisée par cette dernière pour les années 2002 et 2003. * : valeurs obtenues en 2003 après retrait des pixels de bordure. Unités (sauf RRMSE) : t ha^{-1}.

En 2002, sur le dispositif de fertilisation azotée en carré latin, l'approche par cartographie aboutit à des simulations qui s'alignent assez bien par rapport aux observations sur la première bissectrice (Figure IV.3-8). Par ailleurs, la RMSE est beaucoup plus faible qu'en 2000 (1.02 contre 2.27). On constate cependant la même tendance à surestimer la variabilité par rapport à celle qui est observée. Même s'il s'avère que les trois nuages de points correspondant aux trois traitements azotés sont très imbriqués, cet alignement relativement meilleur qu'en 2000 provient essentiellement de ce dispositif qui accroît la variabilité observée. Cela confirme que le modèle réagit correctement à des doses d'azote variées.

L'approche par assimilation donne des simulations assez bien alignées sur une droite parallèle à la première bissectrice, même s'il existe une relative dispersion. La part de la

dispersion sur l'erreur totale est plus faible qu'en 2000 (RMSEu=0.52 et RMSE=2.8 en 2002

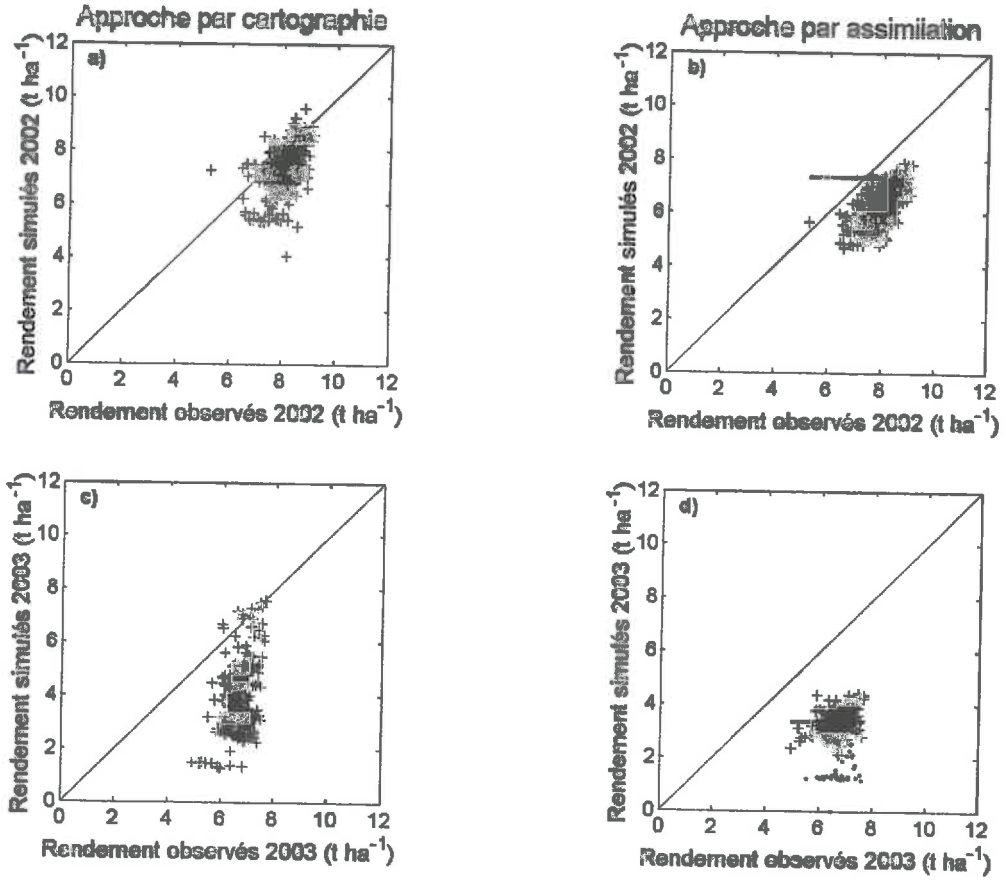


Figure IV.3-8 : Comparaison des simulations de rendements grâce à l'approche par cartographie (a et c) et en utilisant les valeurs des variables d'entrée estimées au cours d'une autre saison culturale par assimilation (b et d) pour les années climatiques 2001-2002 (a et b) et 2002-2003 (c et d). Dans les graphiques concernant l'approche « GLUE », les simulations données par l'information *a priori* sont indiquées par un trait horizontal.

contre RMSEu=0.49 et RMSE=1.24 en 2000), ce qui est également dû à l'introduction de la variabilité liée à la fertilisation azotée. La plus forte valeur de RMSE en 2002 par rapport à 2000 et par cette approche par rapport à l'approche par cartographie provient essentiellement du biais. En effet, si les simulations issues de l'information *a priori* sont en moyenne assez bonnes (RMSE=0.91), les simulations issues des variables estimées par assimilation sont très sous-estimées. Il est malheureusement impossible de porter un diagnostic sur cet état de fait : faute d'observations sur *LAI*, on ne peut savoir (i) si les valeurs des variables d'entrées conduisent à sous-estimer *LAI* (auquel cas, cela signifierait que les valeurs obtenues pour 1999-2000 ne sont pas pertinentes pour 2001-2002 en terme de valeur absolue) ou (ii) si *LAI* est bien simulé mais que, comme observé sur l'autre parcelle en 2000-2001, on peut bien simuler *LAI* tout en sous-estimant la matière sèche et le rendement (auquel cas, il y aurait un

problème dans la modélisation de ces dernières variables). Néanmoins, cette capacité de l'assimilation à introduire une variabilité spatiale grâce à l'utilisation de variables d'entrée estimées au cours d'une autre année de culture suggère d'utiliser ce genre de technique uniquement pour le spatialiser (*cf. infra*). Ce biais peut également être lié au changement de forme de l'engrais : en 2000, certaines variables d'entrée (comme *NORG* ou les paramètres plantes) ont pu être estimées de façon à compenser une sous-estimation par le modèle des pertes d'engrais par volatilisation. En 2002, où l'engrais a une plus forte efficience, ces variables conduisent donc à sous-estimer le rendement (la version 5 de STICS simule les différences de volatilisation des engrais mais le formalisme n'a pas encore été pleinement validé).

Pour ce qui est de l'année climatique 2002-2003 et de la parcelle 2, le modèle réalise, pour les deux approches, des simulations largement sous-estimées. Alors que la moyenne à la parcelle des rendements est de 6.7 t ha^{-1}, l'approche par cartographie l'estime à 3.8 t ha^{-1} et l'approche basée sur l'assimilation à 3.1 t ha^{-1}. Les simulations de 2001 étaient déjà sous-estimées, montrant une trop grande sévérité des stress simulés. L'année 2003 ayant été caractérisée par une forte sécheresse, cela ne fait qu'accroître ce phénomène. Par ailleurs, l'utilisation de l'année climatique 1975-1976 également marquée par une sécheresse importante dans le Chapitre III avait montré la tendance du modèle à surestimer l'effet du stress hydrique. En revanche, les teneurs en protéines atteignaient des valeurs très largement au-dessus du raisonnable, ce qui montre bien que l'absorption d'azote n'était pas limitée et que le stress hydrique, et non azoté, est en cause. Ceci dit, on constate que l'approche basée sur l'assimilation permet encore de relativement bien reproduire la variabilité spatiale : les points qui se fourvoient hors du nuage correspondent à ceux qui le faisaient déjà sur l'année 2000-2001 et qui se trouvent en bordure de parcelle. L'approche par assimilation obtient des RMSEu plus faibles que l'approche par cartographie, surtout après avoir supprimé ces points (RMSEu=0.40 contre 1.25). Ceci confirme la capacité de cette méthode à spatialiser les simulations du modèle, bien qu'elles restent affectées d'un biais très important (RMSEs=3.35). La remarque faite sur le changement de forme de l'engrais s'applique ici également.

IV.3.6 Evaluation de la qualité de la simulation de la variabilité spatiale

La Figure IV.3-9 représente les cartes de rendement simulé de 2000 à 2003 grâce aux deux approches. Mise à part en 2002, il n'est pas aisé de trouver des structures communes

entre ces cartes et celles représentant les observations (Figure IV.2-6). Les cartes de rendement d'une même parcelle et d'une même approche se ressemblent, celles obtenues par assimilation présentent évidemment de fortes similitudes avec celles de *LAI* et *QN*.

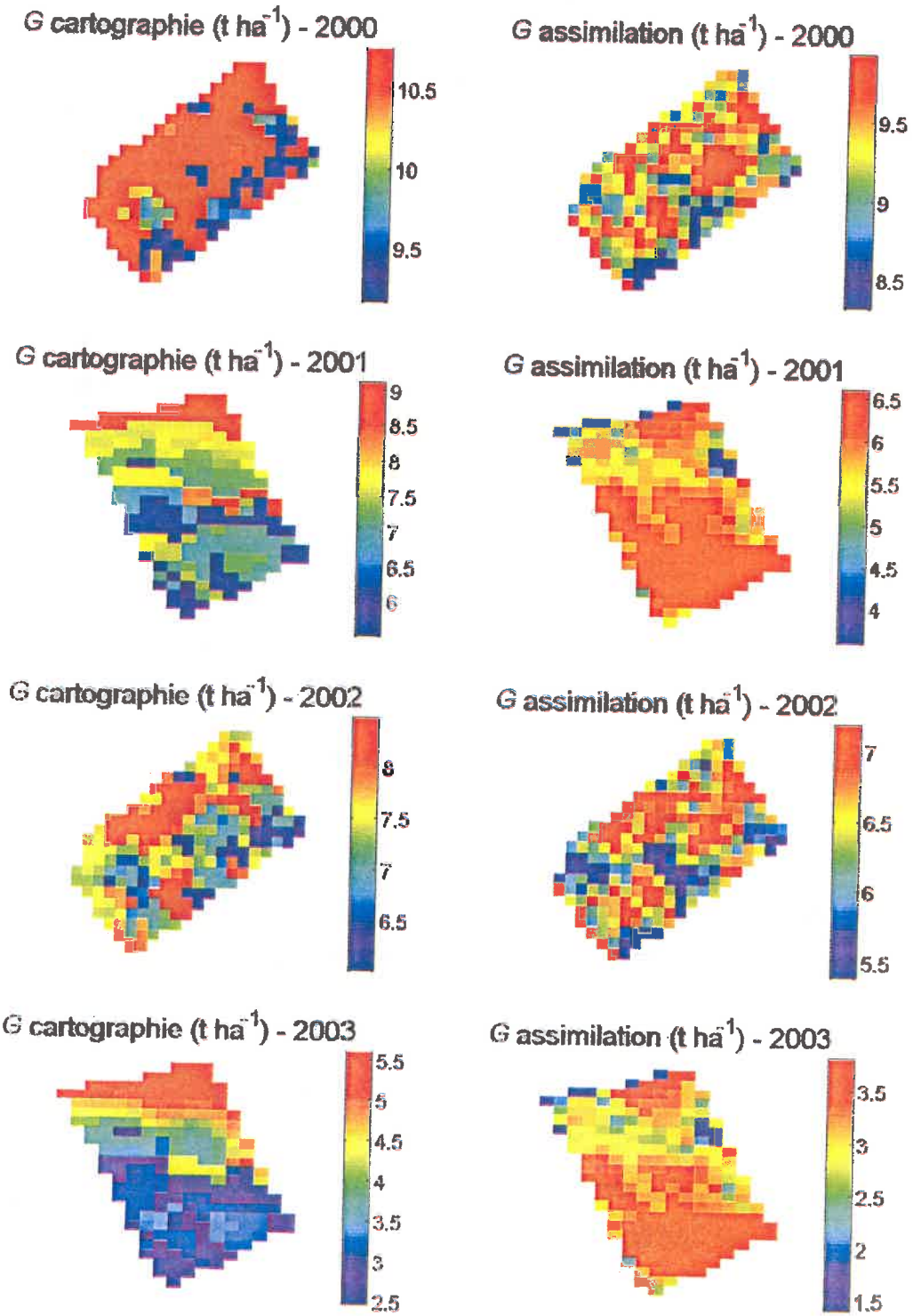


Figure IV.3-9 : Cartes de rendement (*G*) simulé grâce à l'approche par cartographie (à gauche) et pr assimilation (à droite) de 2000 à 2003.

Ce paragraphe est dévolu à une étude plus particulière de la façon dont le modèle simule la variabilité spatiale du rendement. Nous nous limiterons aux années 1999-2000 et 2000-2001. Le principe est de comparer les semi-variogrammes des valeurs observées et simulées, et d'étudier le comportement des variogrammes croisés des valeurs observées et simulées. Une randomisation permet en outre de déterminer si les simulations données par le modèle sont meilleures que des tirages aléatoires.

Un semi-variogramme représente la distribution spatiale de la variable étudiée. Pour une variable Z, il se calcule de façon expérimentale comme suit (Chilès & Delfiner, 1999) :

$$\gamma_{ZZ}(h)=\frac{1}{2}\frac{1}{N(h)}\sum_{\|x-y\|\approx h}\left[(Z(x)-Z(y))^2\right] \qquad \text{Équation IV.3-2}$$

où $\gamma_{ZZ}(h)$ représente la valeur du semi-variogramme calculée pour la classe de distance h, $N(h)$, le nombre de couples de points séparés d'une distance faisant partie de la classe h, $Z(x)$ et $Z(y)$ représentent respectivement les valeurs de la variable Z au point x et au point y séparés d'une distance appartenant à la classe de distance h.

Le semi-variogramme est caractérisé principalement par sa portée et son plateau. La portée est la distance à partir de laquelle il n'y a plus de corrélation spatiale entre $Z(x)$ et $Z(h)$. Il représente la structure spatiale. Le plateau traduit la variabilité spatiale de Z.

La notion de variogramme croisé généralise celle de semi-variogramme. Un variogramme croisé représente la covariance entre les changements de valeur de deux variables dans l'espace. Dans notre cas, il est calculé entre une variable observée (qui peut correspondre au rendement ou à la teneur en protéines), notée O et une variable simulée, notée S. Le variogramme croisé se calcule de façon expérimentale selon l'équation suivante(Chilès & Delfiner, 1999) :

$$\gamma_{os}(h)=\frac{1}{2}\frac{1}{N(h)}\sum_{\|x-y\|\approx h}\left[(O(x)-O(y)).(S(x)-S(y))\right] \qquad \text{Équation IV.3-3}$$

où $\gamma_{os}(h)$ représente la valeur du variogramme croisé de O et S calculée pour la classe de distance h, $N(h)$, le nombre de couples de points séparés d'une distance faisant partie de la classe h, $O(x)$, $O(y)$, $S(x)$ et $S(y)$ représentent respectivement les valeurs de la variable observée au point x et au point y séparés d'une distance appartenant à la classe de distance h et les valeurs de la variable simulée au point x et au point y. Le semi-variogramme γ_{oo} des valeurs observées peut se calculer en utilisant cette même formule en remplaçant S par O et

inversement pour γ_{ss}. Par définition, le variogramme croisé γ_{os} est encadré pour toute classe h de la façon suivante (Chilès & Delfiner, 1999) :

$$|\gamma_{os}(h)| \leq \sqrt{\gamma_{oo}(h).\gamma_{ss}(h)}$$ Équation IV.3-4

Plus le variogramme croisé est proche de l'enveloppe supérieure donnée par $\sqrt{\gamma_{oo}(h).\gamma_{ss}(h)}$, plus il existe une cohérence spatiale entre O et S. Si le variogramme croisé est négatif, cela signifie que O et S évoluent spatialement de façon opposée (quand une augmente, l'autre diminue). Afin d'évaluer la qualité de la simulation du point de vue spatial, on peut utiliser une technique proposée par Peyrard *et al.* (accepté) et utilisée par Baret *et al.* (soumis). On calcule p-1 variogrammes croisés 'aléatoires' $\gamma_{os}^{*}(h)$ en permutant aléatoirement dans l'espace les valeurs de S. Pour tester l'hypothèse que la représentation spatiale du rendement par le modèle est aléatoire avec une erreur de première espèce de α%, on regarde où se situe le variogramme croisé $\gamma_{os}(h)$ par rapport aux $p.\alpha/2$ plus forts et plus faibles variogrammes croisés aléatoires $\gamma_{os}^{*}(h)$ à l'ensemble desquels on rajoute $\gamma_{os}(h)$. Si $\gamma_{os}(h)$ se trouve en dehors de cette limite, on peut rejeter l'hypothèse que la représentation spatiale de la variable par le modèle est aléatoire. Si $\gamma_{os}(h)$ se situe dans cette enveloppe, cela signifie qu'il n'existe aucune structure spatiale entre O et S.

Il est aussi intéressant de comparer les portées des semi-variogrammes de S et de O : si leur portée n'est pas similaire, c'est que S et O n'ont pas la même structure spatiale. La valeur du plateau est également importante à comparer et traduit les différences de variabilité entre S et O.

En résumé, les clefs d'analyse de la Figure IV.3-10 sont les suivantes :

- le variogramme croisé $\gamma_{os}(h)$ doit être prêt de l'enveloppe supérieure $\sqrt{\gamma_{oo}(h).\gamma_{ss}(h)}$ et hors de la limite formée par les $\alpha/2$ et 100-$\alpha/2$ centiles des variogrammes croisés aléatoires $\{\gamma_{os}^{*}(h)- \gamma_{os}(h)\}$;

- le semi-variogramme des valeurs simulées, $\gamma_{ss}(h)$, doit avoir une portée et un plateau proches de ceux du semi-variogramme des valeurs observées, $\gamma_{oo}(h)$.

La Figure IV.3-10 donne ces informations pour les rendements de 2000 et 2001.

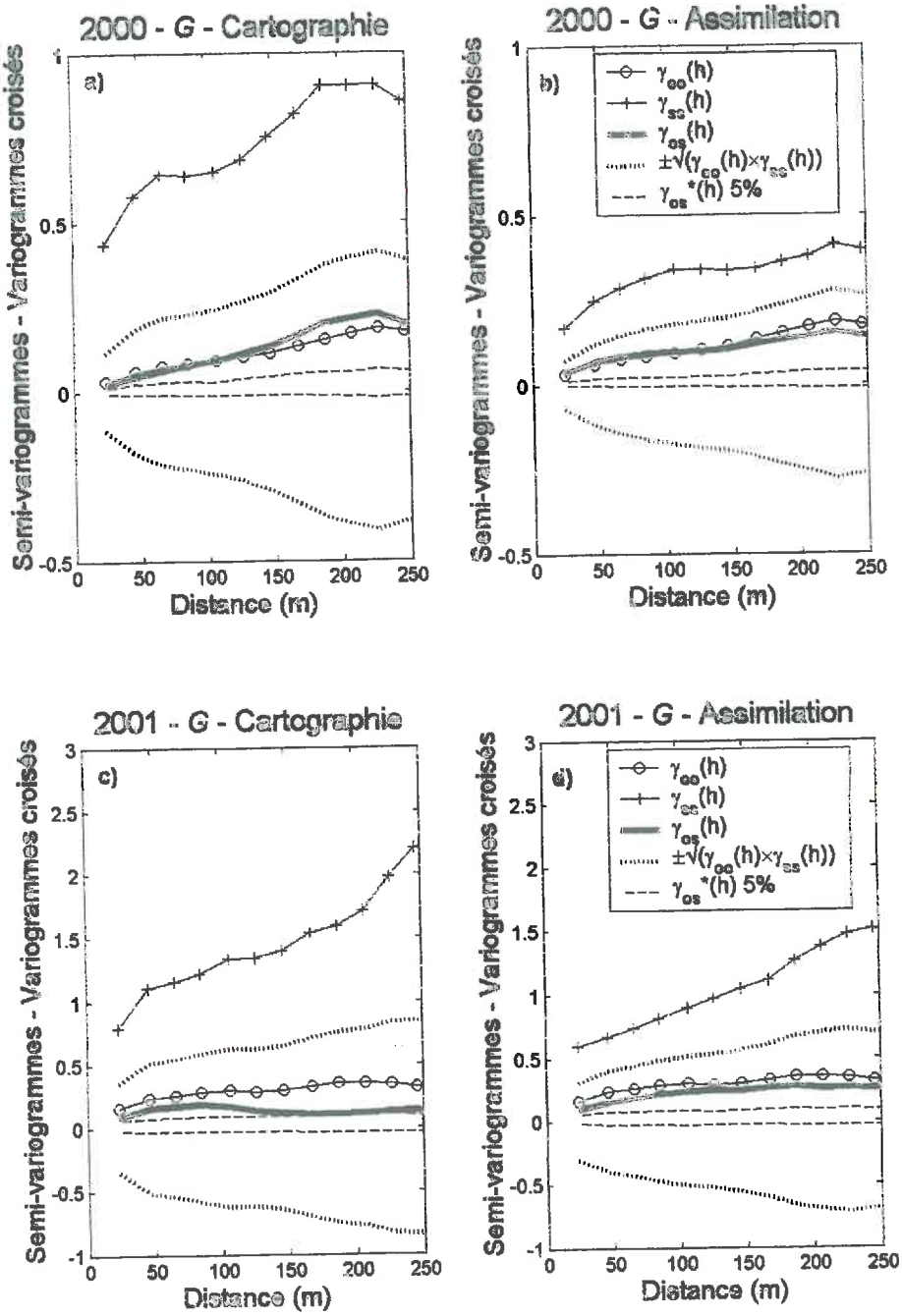


Figure IV.3-10 : Semi-variogramme du rendement observé, $\gamma_{oo}(h)$, du rendement simulé, $\gamma_{ss}(h)$, variogramme croisé des rendements observé et simulé, $\gamma_{os}(h)$, enveloppe $\sqrt{\gamma_{oo}(h).\gamma_{ss}(h)}$ du variogramme croisé, et 5ème et 95ème centiles des variogrammes croisés aléatoires, $\gamma_{os}^*(h)$. a) Rendement simulé grâce à l'approche par cartographie en 2000 ; b) grâce à l'approche par assimilation en 2000 ; c) grâce à l'approche par cartographie en 2001 ; b) grâce à l'approche par assimilation en 2001.

On constate que toutes les variables représentées présentent une variabilité spatiale simulée (donnée par la valeur du plateau du semi-variogramme $\gamma_{ss}(h)$) beaucoup trop forte, ce qui était déjà visible sur les graphiques simulés-observés présentés ci-dessus. Quant aux portées des semi-variogrammes observés et simulés, elles sont rarement proches, ce qui

indique que la structure spatiale des rendements sur la parcelle est assez mal reproduite par modélisation.

Le semi-variogramme du rendement simulé en 2000 se situe à peu près entre l'enveloppe supérieure $\sqrt{\gamma_{oo}(h).\gamma_{ss}(h)}$ et le 95ème centile des variogrammes croisés aléatoires, ce qui traduit une qualité moyenne de la restitution de la variabilité spatiale. En 2001, la qualité de la simulation de la variabilité spatiale par l'approche cartographique à longue distance (à partir de 150 m) n'est pas meilleure que les tirages aléatoires. La simulation basée sur l'assimilation donne en revanche des résultats du même ordre de qualité de ce point de vue que sur l'autre parcelle en 2000.

L'observation de ce genre de graphiques concernant les années 2002 et 2003 montre que pour la première de ces années, la spatialisation des simulations est assez bonne, ce qui est, comme on l'a vu, largement attribuable aux traitements azotés. En 2003 en revanche elle est assez mauvaise pour tous les types d'approches, sauf si on supprime les points dont on a vu qu'ils sortaient systématiquement du nuage de points dans l'approche par assimilation et qui correspondent aux pixels situés en bordure de parcelle. La suppression de ces points confirme que la variabilité des rendements est bien restituée par l'approche basée sur l'assimilation.

Ce type d'analyse, facile à mettre en œuvre, pourrait être employé dans des prolongations de ce travail afin de caractériser la qualité de la spatialisation des simulations.

IV.3.7 Discussion sur les assimilations de 2000 et 2001

Nous revenons ici sur les différents problèmes qui ont pu perturber les résultats de l'assimilation. Sur la parcelle 1 en 1999-2000, l'assimilation conduit à d'assez bons résultats sur la simulation du rendement, tandis que sur le jeu de 2001-2002, les simulations sont très biaisées sur le rendement, sans qu'on puisse savoir si *LAI* est bien simulé, faute d'observations. Sur la parcelle 2, l'assimilation permet de bien simuler *LAI* l'année de l'assimilation, mais le rendement est fortement biaisé, quoique la variabilité spatiale soit assez bien reproduite. L'utilisation sur le jeu de 2002-2003 des valeurs de variables d'entrée retrouvées conduit à un biais encore plus fort, montrant que les stress hydrique et azoté sont très exagérément simulés.

Pour expliquer ces constations, une première hypothèse consisterait à incriminer l'information *a priori*. Mais parmi les simulations réalisées d'après cette information, il s'en

trouve qui donnent des rendements très proches de ceux observés. Le problème vient donc de la capacité de trouver des jeux de variables d'entrée qui permettent de bien simuler à la fois *LAI* et *QN*, les variables servant à l'assimilation, et les variables d'intérêt final, comme le rendement.

D'autres explications consistent à partir du constat qu'on simule correctement *LAI* et *QN* mais pas le rendement (et les teneurs en protéines). Pour l'expliquer, on peut remettre en cause (i) la structure du modèle ; (ii) la façon d'estimer les variables d'état ; (iii) les valeurs de *LAI* et *QN* observées qui servent à réaliser l'assimilation. Après avoir examiné ces trois points, cette partie dégage les conséquences qui doivent être tirées de cette expérience.

IV.3.7.1 Analyse du problème

Après mon septième apéritif, je pensais qu'il n'y a ni réel, ni irréel, ni mensonge. (...) C'est bien malin de philosopher sans avoir appris à philosopher et après sept apéritifs.
Ionesco, Le Solitaire

IV.3.7.1.1 Erreur de structure du modèle

Selon cette hypothèse, si *LAI* et *QN* sont bien simulées alors que le rendement ne l'est pas, c'est à cause de l'écriture même du modèle. Cette hypothèse peut être soutenue par le constat fait dans le Chapitre II que STICS a tendance à surestimer les valeurs de *LAI* au maximum du développement, sans que cela induise une surestimation du rendement. On sait en effet que plus que la valeur de ce plateau, ce sont les phases de début et de fin de cycle de *LAI* qui déterminent la matière sèche et le rendement. Ainsi, si l'on observe des LAI au niveau du plateau plus faibles que ceux simulés, l'assimilation va chercher à les diminuer et le prix en sera une avancée de la sénescence et donc une diminution des rendements simulés.

On peut trouver une explication à cette erreur de structure dans la simulation des stress. L'effet des stress sur *LAI* est représenté par la plus petite valeur des stress azoté et hydrique, alors que l'effet des stress sur la matière sèche est le résultat du produit des deux types de stress (*cf.* description de STICS, Chapitre II). Ainsi, lorsque *LAI* est utilisé dans l'assimilation, le processus cherche à le réduire en créant des stress par les variables sol : on introduit les deux types de stress mais un seul va contribuer à une date donnée à diminuer *LAI*. Par contre, les deux stress s'appliqueront de façon combinée sur la matière sèche et par conséquent sur le rendement, ce qui conduit à exagérer l'effet des stress sur cette dernière variable.

Il faut également rajouter dans les erreurs de structure du modèle des aspects déjà discutés : la non prise en compte de certains facteurs limitant par le modèle et les instabilités numériques.

IV.3.7.1.2 Mauvaise estimation des variables d'entrée

L'estimation des variables d'entrée caractéristiques du sol et de la plante par l'assimilation serait la cause du biais observé sur le rendement, ce qui peut être imputé à différents aspects de la méthode :

- Les valeurs attribuées aux variables d'entrée :

La première raison est la différence de sensibilité de *LAI* et *QN* d'une part et du rendement d'autre part aux variables d'entrée. En effet, pour modifier les variables servant à l'assimilation, on peut être amené à modifier une variable d'état dans des proportions qui affectent beaucoup plus le rendement : ceci rejoint le problème de la structure du modèle.

Par ailleurs, de nombreuses variables d'entrée ne sont pas ré-estimées, soit parce qu'elles n'ont pas été versées par l'utilisateur dans le pool des variables susceptibles de l'être, soit que la distribution *a posteriori* obtenue par GLUE correspond à la distribution *a priori*. Dans ce cas, la valeur de l'information *a priori* de ces variables peut conditionner la qualité des simulations du rendement si celui-ci est sensible à ces variables. Ceci peut être le cas des *HCCF(h)* : les figures III.3-10 et 11 montrent qu'elles sont beaucoup plus influentes sur le rendement que sur *LAI* et *QN* et on vient de voir qu'elles avaient été effectivement peu ré-estimées en 2001. Ceci montre l'importance de l'information *a priori* et peut expliquer une absence de cohérence entre les valeurs simulées de rendement et de *LAI*/*QN*.

On ne peut pas réellement imputer au nombre important de variables d'entrée que l'on cherche à estimer la mauvaise qualité relative des résultats. En effet, la méthode GLUE opère, comme on l'a vu, un tri parmi les variables qu'on se propose d'estimer. Celles auxquelles le modèle n'est pas sensible dans ces conditions sont laissées à leur valeur donnée par l'information *a priori*. Ceci revient, implicitement, à leur attribuer une valeur et à n'estimer que quelques variables. L'intérêt d'en intégrer un nombre relativement important dans le système est de conserver une souplesse qui permet de s'adapter aux changements de conditions, pédo-climatiques par exemple, vu qu'on a montré que la sensibilité du modèle aux variables d'entrée dépend de ces conditions. Ceci affranchit dans une certaine mesure de l'analyse de sensibilité.

- La position des dates d'observation par rapport aux variables d'entrée estimées :

Nous avons vu dans le Chapitre III sur des données simulées que les dates les plus importantes pour réaliser l'assimilation se situent en fin de cycle. Les jeux de données de 2000 et de 2001 en donnent une illustration concrète : en 2000 les observations vont jusqu'à fin juin, alors qu'en 2001 elles s'arrêtent en mai. *durvieF* par exemple a été utilisé en 2001 'pour' diminuer les valeurs de *LAI* observées, mais comme il n'y a pas de dates d'observation après le 31/05, cette diminution a pu être trop forte, avec comme conséquence une sous-estimation systématique du rendement. Cette même année, *LAI* et *QN* ne sont que peu sensibles à *OBSTARAC* qui est de ce fait laissée à sa valeur donnée par l'IAP (*cf.* Figure IV.3-4 et explications). La valeur de l'information *a priori*, centrée autour de 100 cm, permet de simuler correctement les observations de *LAI/QN* jusqu'à fin mai. Par contre, après cette date, cette variable d'entrée prend plus d'importance et sa valeur faible affecte le rendement. On est ici dans un cas où la valeur de l'information *a priori* induit en erreur du fait d'un manque de données observées en fin de cycle.

Dans le Chapitre III, nous avions également montré que le nombre de dates joue peu sur la qualité de l'assimilation en utilisant des jeux de données simulées. Mais il est probable que ce soit moins le cas avec des données réelles : en effet, si des observations sont acquises au cours d'un stress passager, elles pourront avoir un fort impact si elles ne sont pas contrebalancées par de nombreuses autres observations effectuées dans des conditions plus représentatives des conditions générales de croissance du couvert (Launay, 2002).

- Le non respect des corrélations entre les valeurs des variables d'entrée :

Le manque de cohérence des simulations de *LAI* et de rendement mis en évidence par l'étude des résidus serait lié, outre les raisons évoquées ci-dessus, à l'absence de prise en compte des corrélations qui existent entre les variables caractérisant le sol.

IV.3.7.1.3 Mauvaise estimation de *LAI* et *QN*

Les observations de *LAI* et *QN* peuvent elles-mêmes être erronées et estimées trop faibles par exemple. En cherchant à diminuer les valeurs simulées de ces variables, on ne peut qu'aboutir à des valeurs systématiquement sous-estimées du rendement. Un exemple flagrant de cette situation est fourni par les pixels de bordure de parcelle où *LAI* et *QN* sont sous-estimés : les rendements correspondants simulés sont eux-mêmes très faibles.

L'erreur commise sur l'estimation de *LAI* et *QN* peut être un biais : en effet, c'est la valeur absolue de ces variables qui est difficile à obtenir plus que les valeurs relatives. Ainsi, une estimation biaisée des variables observées ne peut que conduire à simuler un rendement lui-même biaisé.

IV.3.7.2 Conséquences à envisager

Afin de prendre en compte le fait que le modèle a naturellement tendance à surestimer *LAI* au plus fort du plateau, il faudrait ne pas utiliser d'observations au cours de cette période et se focaliser sur les dates de début et, surtout, de fin de cycle.

Pour arriver à estimer des valeurs de variables d'entrée qui permettent de simuler des valeurs de *LAI/QN* cohérentes avec celles de rendement, l'idéal serait de prendre en compte des données de rendement même. Ceci ferait perdre tout espoir d'utiliser les mesures de l'année en cours pour préconiser des doses d'azote ou prévoir des rendements. Pour respecter cette cohérence, on pourrait également prendre en compte les corrélations entre variables d'entrée. Des méthodes d'assimilation existent qui permettent d'évaluer les corrélations entre les variables d'entrée estimées, mais il faut posséder une information *a priori* sur ces corrélations qui sont également spécifiques du lieu (*cf.* Chapitre II et bibliographie sur les fonctions de pédotransfert), ce qui complique le système. Le meilleur moyen de respecter ces corrélations reste peut-être d'utiliser le rendement dans le processus d'assimilation.

Une autre solution consisterait à n'utiliser les observations de *LAI/QN* que pour réaliser la spatialisation des simulations du rendement. Nonobstant le biais ainsi introduit sur cette variable, on pourrait utiliser la méthode pour spatialiser les préconisations d'azote effectuées en moyenne sur la parcelle par des méthodes éprouvées. En poursuivant ce raisonnement, on peut aussi proposer de ne chercher à utiliser que l'information spatiale contenue dans les cartes de *LAI* et de *QN* : il faudrait alors utiliser dans les algorithmes d'assimilation un critère à minimiser qui fît fi du biais entre les simulations de ces variables et les valeurs observées. Ceci aurait la vertu de permettre de s'affranchir d'un biais éventuel sur l'estimation de ces variables par inversion des modèles de transfert radiatif.

IV.4 Mise en œuvre expérimentale de la méthode

Cette partie est l'occasion de faire une synthèse des chapitres II et III, le premier discutant du choix d'un critère agro-environnemental permettant de déterminer des doses optimales, le second développant deux méthodologies de spatialisation du modèle de culture. L'objectif était de mettre en œuvre la méthode de préconisation spatialisée de doses d'azote en conditions réelles sur le site de Chambry au cours de la saison de culture 2002-2003, sur la parcelle 2, pour le troisième apport, c'est-à-dire vers la fin du mois d'avril. A cette époque, seule la caractérisation des propriétés du sol par l'approche pédologique était opérationnelle et c'est donc celle-ci qui a été utilisée. Cependant, au vu des résultats obtenus sur l'année climatique 2000-2001 sur la même parcelle, il a été procédé à une réestimation d'une des variables d'entrée du sol. Les paragraphes qui suivent présentent successivement cet ajustement, la méthode de simulation utilisée, l'utilisation du critère et enfin le passage de la carte préconisée à la carte réellement épandue. Enfin, une évaluation rapide de l'intérêt de réaliser une fertilisation azotée modulée sera effectuée, car une moitié de parcelle a effectivement reçue des doses d'azote spatialement variables tandis que l'autre a reçu une dose homogène.

IV.4.1 Caractérisation du sol utilisée

Dans cette étude, toutes les simulations ont été effectuées sur la base de 1066 pixels de 10×10 m^2. Les cartes de teneurs en eau et en azote du sol au semis sont les mêmes que celles décrites dans le Chapitre III. Cependant, il a été décidé, afin de prendre en compte des résultats préalables d'évaluation des simulations sur la même parcelle au cours de l'année climatique 2000-2001, de réestimer la valeur d'*OBSTARAC* par Unité Cartographique de Sol. En effet, cette variable a, typiquement, une signification différente selon que l'on cherche à décrire le sol ou les interactions sol-plante. On a donc jugé qu'elle était vraisemblablement mal caractérisée pour le modèle par l'approche cartographique, notamment d'après les faibles valeurs prises (nombreuses valeurs inférieures à 80 cm). Une simple optimisation de cette variable d'entrée sur des simulations effectuées à l'échelle des UCS a été effectuée par rapport aux rendements. Il était très vraisemblable que l'adaptation d'une seule variable d'entrée par rapport à une seule variable de sortie n'était qu'un pis-aller et que les valeurs obtenues ne conviendraient vraisemblablement pas pour une autre année climatique : ceci sera évalué dans ce paragraphe.

IV.4.1.1 Adaptation des valeurs d'*OBSTARAC* sur le jeu de données de 2001

Pour chaque UCS, le modèle a donc été considéré comme une fonction qui à la variable *OBSTARAC* associait le rendement. *OBSTARAC* a ainsi été ajusté à une valeur qui permettait de simuler le rendement obtenu pour l'UCS considérée d'après la carte des rendements. Vu la réponse du modèle à cette variable (*cf. supra*), cette opération fonctionne en général parfaitement bien. Cependant, pour quatre UCS, aucune valeur d'*OBSTARAC* ne permettait d'atteindre le rendement observé, car les valeurs attribuées aux autres variables de sol étaient trop limitantes. Pour ces UCS, les valeurs d'origine ont été conservées.

IV.4.1.2 Validation *a posteriori* de cette adaptation sur les résultats de 2003

Afin d'évaluer les répercussions de la réestimation d'*OBSTARAC* grâce aux données de 2001 sur les simulations de 2003, nous donnons ici la comparaison des simulations effectuées avec les valeurs d'*OBSTARAC* estimées par l'approche pédologique et les valeurs réestimées comme décrit ci-dessus.

On constate que les valeurs d'*OBSTARAC* réestimées pour la saison 2000-2001 conduisent à de meilleures simulations que celles qui avaient été initialement établies. Néanmoins, cette amélioration est assez faible, ce qui était attendu : l'amélioration vient de l'augmentation de la réserve utile induite par l'augmentation générale de valeurs d'*OBSTARAC*, ce qui se traduit par une augmentation des rendements moyens. Mais ils demeurent toujours trop faibles, montrant la trop forte sensibilité de l'ensemble {modèle-caractérisation du sol} au déficit hydrique, particulièrement important en 2003. Seules les valeurs d'*OBSTARAC* du nord de la parcelle avaient été diminuées sur la foi des observations de la saison 2000-2001, et c'est précisément dans cette zone que les rendements maximums n'ont pas été atteints. Ces données de « validation » sont indiquées à l'échelle des UCS car c'est à cette échelle que les valeurs d'*OBSTARAC* ont été réestimées.

Il n'en demeure pas moins que la carte de préconisation a été établie sur cette base de simulation.

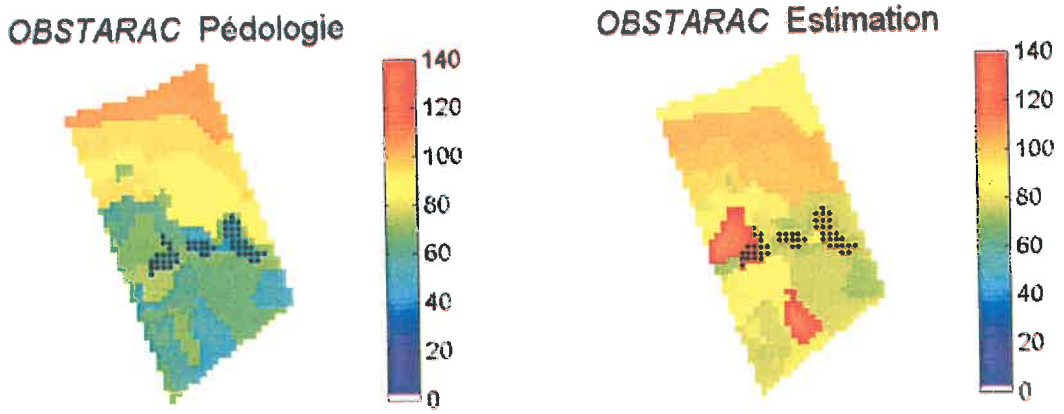


Figure IV.4-1 : Comparaison des valeurs de la variable d'entrée sol *OBSTARAC* établies par l'approche pédologique et estimées par ajustement des valeurs de rendements au niveau des UCS pour l'année climatique 2000-2001. Les pixels marqués d'un point correspondent à des zones où l'ajustement n'a pas été possible à cause des valeurs trop limitantes des autres variables d'entrée.

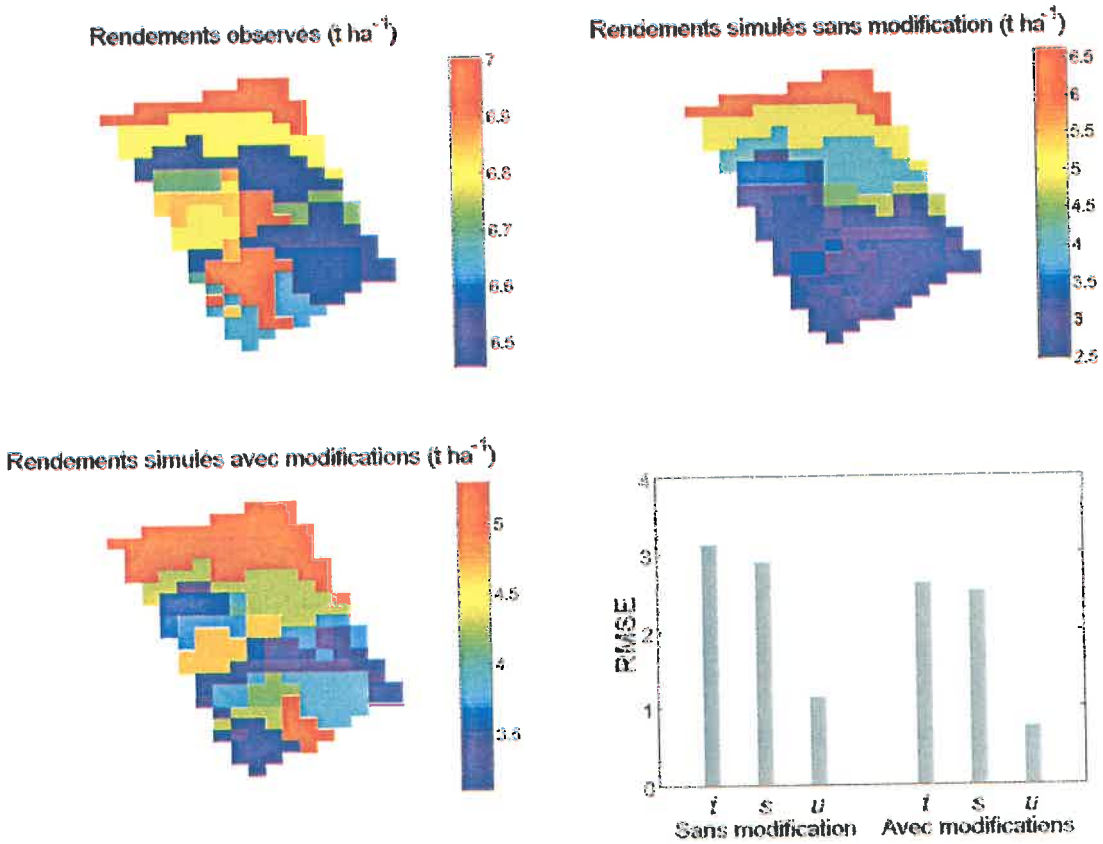


Figure IV.4-2 : Comparaison de la qualité des simulations effectuées pour la saison culturale 2002-2003 sur la base des données sol déterminées par l'approche pédologique (« sans modification ») et en utilisant les valeurs d'*OBSTARAC* réestimées par rapport au rendement sur la saison 2000-2001 (« avec modifications »). Pour chacune des deux façons de faire, les RMSE totales (*t*), RMSEs (*s*) et RMSEu (*u*) sont indiquées.

IV.4.2 Réalisation des simulations croisant scénarios techniques et climatiques

Afin de choisir une dose d'azote pour le troisième apport en prenant en compte l'incertitude climatique entre la date de décision et la date de la récolte, des simulations croisant scénarios climatiques et scénarios techniques ont été effectuées. Afin de limiter le temps de calcul, seul 10 climats et 9 itinéraires techniques ont été utilisés.

En ce qui concerne le climat, les données météorologiques réelles ont été utilisées de la date de semis jusqu'au 31 mars, figurant la date de décision. Entre le 1er avril et la récolte, 10 séries climatiques ont été utilisées, choisies principalement dans les années 1990 (6 sur 10) pour prendre en compte la tendance du climat à évoluer, et complétées par des données des années 1970 et 1980 afin d'augmenter la variabilité des climats.

Les 9 scénarios techniques testés ne différaient que par la quantité d'azote épandue au troisième apport. Les premier et deuxième apports ont été fixés aux doses réellement apportées, 40 kg ha^{-1} au 11 mars et 60 kg ha^{-1} au 26 mars et le quatrième a été supposé constant sur la parcelle et fixé à la dose prévue de 40 kg ha^{-1} au 20 mai. Cet apport a finalement été fait le 12 mai. La date de l'apport modulé a été *a priori* fixée au 24 avril et a effectivement eu lieu le 23.

IV.4.3 Choix et application du critère agro-environnemental

Au moment de l'établissement de la carte, le travail sur le choix du critère agro-environnemental décrit dans le Chapitre II n'était pas terminé. L'élaboration de cette carte a du reste permis d'éclairer cette réflexion. Il ne s'agit pas pour nous de revenir ici sur la question de ce choix mais, brièvement, nous pouvons dire qu'à cette occasion il a été envisagé d'utiliser la quantité d'azote lessivé au cours de l'hiver suivant la récolte sous hypothèse de sol nu comme variable environnementale. Il s'est avéré que cette dernière, tout comme les reliquats, était beaucoup plus soumise aux propriétés du sol qu'à la fertilisation azotée dans le modèle. En outre, ces simulations ne pouvaient en aucune manière être validées. Même si l'utilisation de cette variable pour la décision de la fertilisation azotée serait pertinente du point de vue environnemental, c'est le bilan d'azote qui a été choisi comme variable environnementale et a été pris en compte par un seuil comme décrit plus haut au Chapitre II. La marge brute, autre composante du critère agro-environnemental rappelons-le, a été calculée selon les prix pratiqués par la coopérative de Chambry.

Des cartes de préconisation ont été établies grâce aux moyennes des simulations réalisées avec les différents climats en utilisant différentes valeurs de seuil sur le bilan d'azote. Le critère finalement choisi fut la dose qui maximisait la marge moyenne tout en conduisant à un bilan d'azote moyen inférieur à un seuil de 50 kg ha^{-1}, ce qui constitue déjà une contrainte importante.

Enfin, une étude ultérieure à celle-ci a montré qu'il serait plus pertinent de calculer les doses optimales d'azote de façon à ce que le critère soit respecté non pas pixel par pixel, mais à l'échelle de la parcelle entière (Houlès *et al.*, 2003).

IV.4.4 Passage de la carte préconisée à la carte réellement appliquée

IV.4.4.1 Comparaison de la carte consigne et de la carte d'épandage effectué

La carte de préconisation a été établie sur des pixels de 10×10 m^2. La fertilisation spatialement modulée a été effectuée par le CEMAGREF avec un épandeur centrifuge d'engrais solide, sous forme de granulés. La résolution de ce type d'appareil est de 24 m latéralement, ce qui correspond à la distance de projection couverte par l'épandeur ; il n'est pas aussi aisé de définir une résolution dans le sens d'avancement du tracteur, vu qu'elle dépend de l'inertie du système à répondre à un changement de consigne, de l'écart entre les deux consignes et de la vitesse d'avancement. Elle est évaluée de l'ordre de la dizaine de mètres (Piron, comm. pers.). La carte préconisée a donc été dégradée afin de s'adapter à ces contraintes. En outre, afin d'évaluer expérimentalement l'intérêt de réaliser une fertilisation modulée, la moitié ouest de la parcelle a été fertilisée à une dose consigne égale à la moyenne des doses apportées sur la zone est. Enfin, un petit secteur sis au nord-est, siège d'essais de doses d'azote, a été l'objet d'un traitement particulier. La Figure IV.4-3 permet de comparer les cartes de préconisation et d'épandage réel, établie par des mesures de flux d'engrais en sortie d'appareil. Il est à noter que cette carte est elle-même entachée d'erreurs (liées essentiellement à la différence entre la mesure du débit et la mesure de la répartition de l'engrais au sol), mais c'est celle-ci qui servira d'entrée au modèle STICS. On constate que sur la zone modulée, les différentes zones sont globalement bien traduites, bien qu'elles soient lissées, notamment dans le quart sud-est, ou beaucoup plus perturbées (par les débuts de ligne et les fourrières), comme dans la partie nord. La zone non modulée avait une dose consigne de 37 kg ha^{-1}, correspondant à la moyenne des doses apportées dans la zone modulée. En réalité, les doses épandues sur cette zone ont une moyenne de 35 kg ha^{-1}, ce qui est très proche de la valeur de consigne mais avec un écart-type de 7.5 kg ha^{-1}, ce qui montre une certaine

instabilité des flux d'engrais en sortie de l'appareil même pour une consigne constante. La dispersion que montre le graphique indiquant les doses réellement apportées dans la zone modulée en fonction des doses préconisées peut avoir deux causes : (i) le lissage des valeurs appliquées par rapport à la dose de consigne lié à la résolution de l'appareil et (ii) l'instabilité du flux d'engrais. Mais le biais sur les doses moyennes est très faible (RMSEs=8.7 kg ha^{-1}).

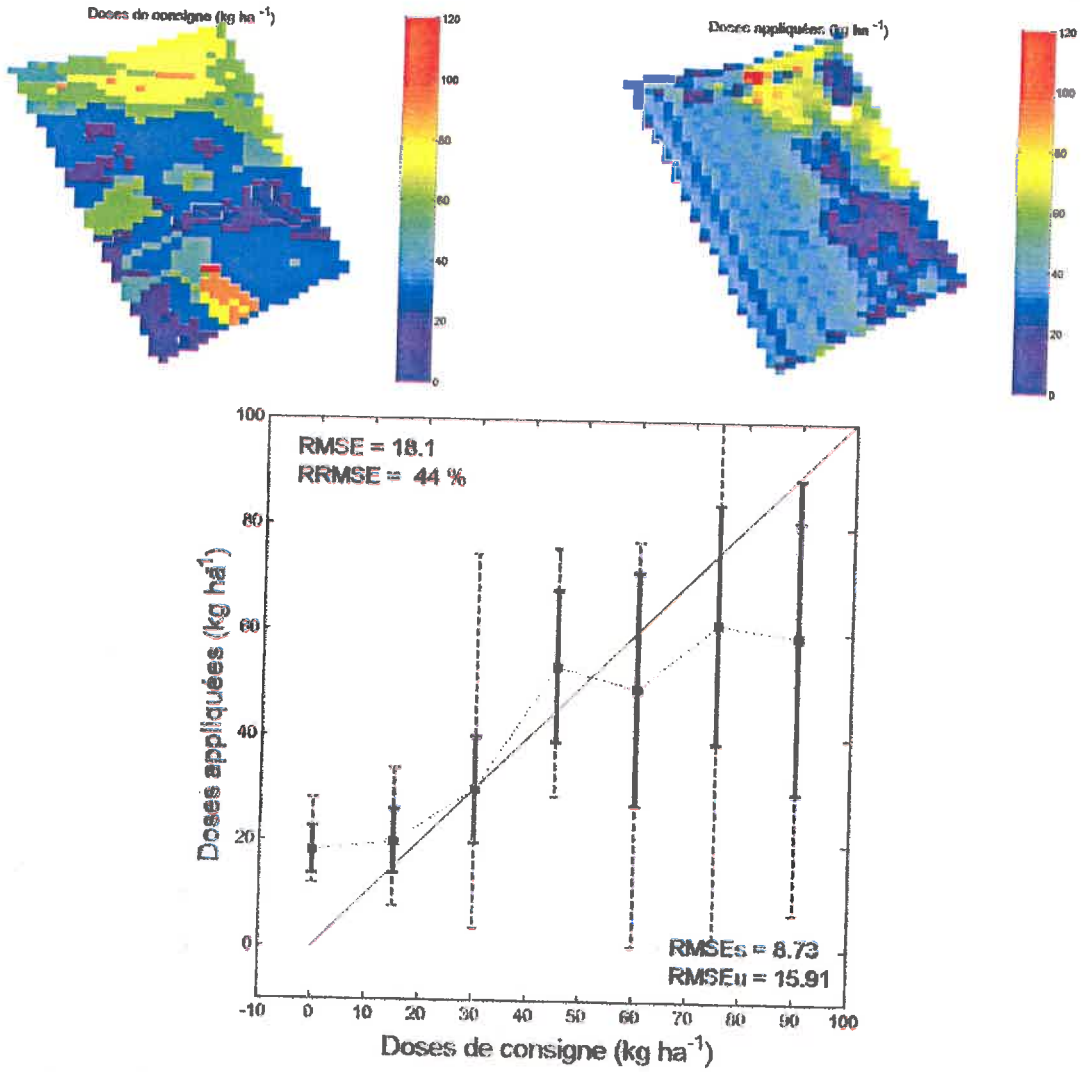


Figure IV.4-3 : En haut à gauche : carte de fertilisation azotée préconisée ; en haut à droite : carte appliquée d'après la mesure CEMAGREF (Piron, comm. pers.). La zone ouest a été fertilisée à une dose égale à la moyenne des apports effectués dans la zone est, de 37 kg ha^{-1}. Une petite zone au nord est correspondait à des essais de doses d'azote et n'a pas été fertilisée. En bas : différence des moyennes des doses consignes et des doses apportées sur la zone modulée. Les barres d'erreur en traits pleins représentent les écarts types ; les barres d'erreur en traits pointillés les valeurs minimales et maximales.

IV.4.4.2 Conséquences à envisager

Les deux contraintes techniques lors de l'épandage mises en évidence ci-dessus, à savoir (i) la résolution relativement grossière de l'appareil et (ii) les différences constatées

entre les doses consignes et les doses mesurées, devraient être prises en compte dans le système permettant de réaliser les préconisations.

D'une part la dose optimale pourrait être raisonnée à l'échelle du « pixel » correspondant à la résolution de l'appareil. Le problème est que ce pixel est bien défini dans la direction perpendiculaire à la direction du tracteur (largeur de l'épandage), mais mal dans la direction parallèle, à cause du temps de réponse du système à un changement de consigne. Le choix de la dose optimale pourrait se faire au niveau de bandes correspondant à la largeur d'épandage qui ne peuvent recevoir qu'une seule dose, plutôt qu'au niveau de pixels de dimension inférieure.

L'inertie du système et les erreurs par rapport à la dose consigne pourraient servir à définir les zones dans le sens parallèle au déplacement de l'appareil, en faisant en sorte d'éviter les transitions brutales.

IV.4.5 Evaluation expérimentale rapide de l'intérêt de moduler la fertilisation azotée à l'échelle de la parcelle

Afin d'évaluer expérimentalement l'intérêt de faire varier géographiquement la dose d'azote, une moitié de la parcelle seulement a fait l'objet d'un apport modulé, tandis que l'autre recevait un apport constant. Nous donnons ici une première approche de cette analyse, consistant simplement à comparer les valeurs obtenues expérimentalement des rendements, teneurs en protéines, marges brutes et bilans d'azote. Ceci ne constitue qu'une analyse frustre car il est nécessaire de prendre en compte l'effet du sol. Cette analyse plus poussée fera l'objet d'un travail au sein du laboratoire de Biométrie de l'INRA d'Avignon.

La Figure IV.4-4 représente les distributions des différentes variables d'intérêts mesurées sur les zones modulées et non modulées de la parcelle. On retrouve le fait que les doses appliquées sur la zone non modulée sont dispersées autour de la dose consigne. Ensuite, un test statistique de Student permet de montrer qu'il existe une différence significative entre les deux zones avec une erreur de première espèce de 10% sur les rendements (bien que très faible) et les teneurs en protéines : cette différence est respectivement de -0.14 t ha^{-1} et de +0.50 % entre la zone à fertilisation modulée et la zone à fertilisation uniforme. Cette augmentation de la teneur en protéine sur la zone modulée est intéressante, mais il est difficile de l'attribuer à la modulation. La différence sur les marges brutes et les bilans d'azote n'est pas jugée significativement différente et est respectivement de -6.6 € ha^{-1} et de +1.4 kg ha^{-1}.

Ceci peut très bien s'expliquer par le fait que la diminution du rendement est compensée par l'augmentation de la teneur en protéines.

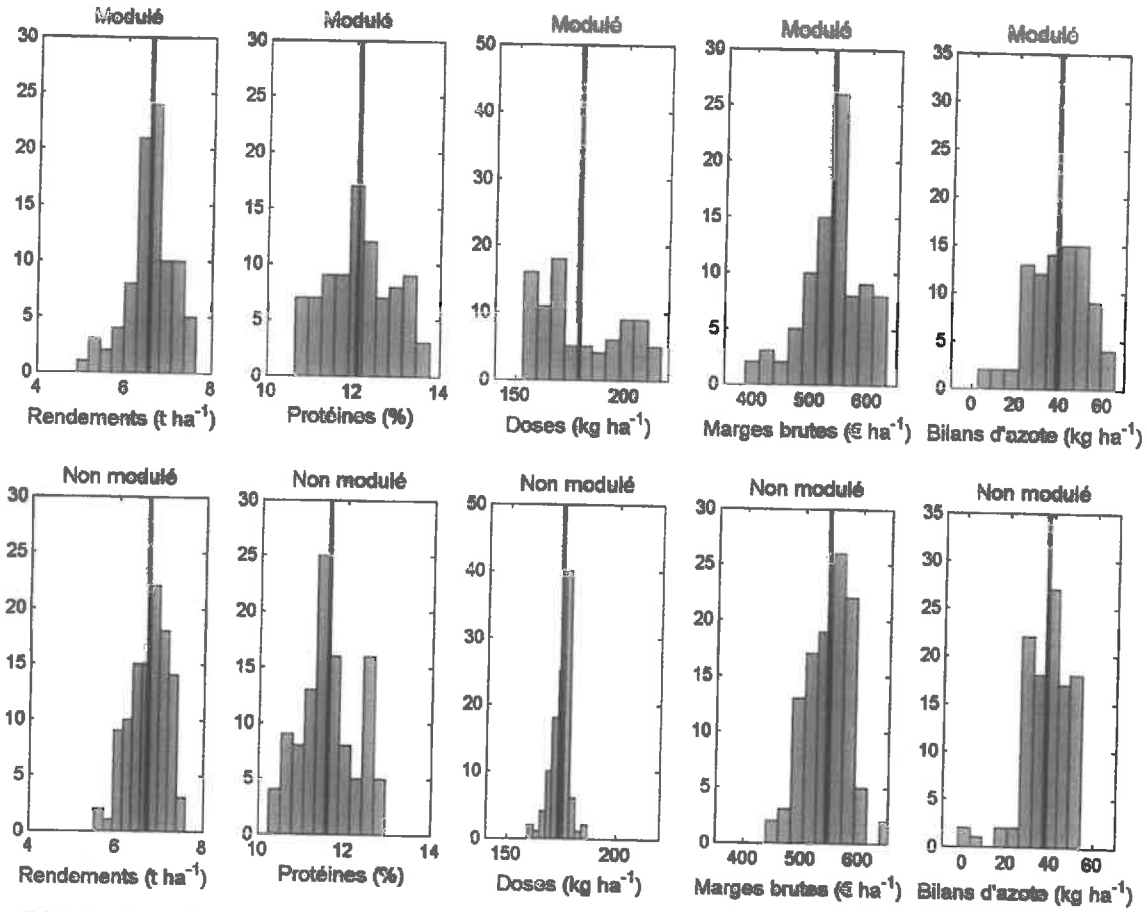


Figure IV.4-4 : Représentation des fréquences de valeurs de rendement, teneur en protéines, doses effectivement apportées, marges brutes et bilans d'azote et de leur moyenne (trait fin) pour la zone de la parcelle ayant reçu un apport modulé et la zone ayant reçu un apport constant.

Bien qu'on ne puisse pas conclure à partir d'une telle analyse sur l'intérêt de la modulation, il semble en première approche assez limité, si ce n'est sur la teneur en protéines. En outre, la mauvaise qualité des simulations qui ont servi à réaliser la préconisation spatialisée laisse penser que ce qui peut être jugé à travers ce jeu de données est seulement l'intérêt d'utiliser l'approche de spatialisation par cartographie pour faire une préconisation spatiale. On conçoit fort bien que l'intérêt de réaliser une modulation soit dépendante du degré d'hétérogénéité du sol, ainsi que de la structure de cette hétérogénéité, qui doit être appréhendée en fonction de la résolution de l'épandeur. Ceci sera évalué dans la suite de ce travail en réalisant une sorte d'analyse de sensibilité par simulation de la différence entre le critère agro-environnemental obtenu en modulant et sans moduler en fonction du degré d'hétérogénéité du sol, exprimé par exemple en terme de rendement potentiel.

IV.5 Conclusion

Cette partie a permis de mettre en évidence plusieurs obstacles méthodologiques dans la possibilité de préconiser des doses d'azote optimales à l'échelle intra-parcellaire en utilisant un modèle de culture. Le premier d'entre eux est la qualité même de la simulation à l'échelle intra-parcellaire des variables impliquées dans le calcul du critère agro-environnemental, même si on sait qu'il n'existe pas de lien direct entre qualité des simulations des variables et qualité des préconisations. Les simulations obtenues par l'approche cartographique sont en général meilleures en moyenne, mais seule l'approche par assimilation permet de modéliser la variabilité intra-parcellaire du rendement d'une façon qui commence à être satisfaisante, même si elle peut être affectée d'un important biais. Ceci rejoint le constat de Wendroth *et al.* (2003) que le rendement est mieux expliqué par les images de télédétection que par les propriétés du sol déterminées par une approche cartographique. Pour autant que notre jeu de données permette d'en juger, cette simulation de la variabilité intra-parcellaire est retrouvée en utilisant pour une année donnée les valeurs des variables d'entrée sol du modèle estimées par assimilation une autre année. Néanmoins, il est fondamental de considérer que le choix, non définitif, qui a été fait de comparer les simulations aux observations sur des pixels de 20×20 m^2 alors que toutes les mesures avaient été réalisées à des échelles différentes et variées n'est pas pour rien dans la dispersion des résultats. Ceci est un point qui mérite réflexion. Les biais obtenus dans l'approche par assimilation semblent provenir d'une différence entre la cohérence des variables {*LAI* et *QN*} et rendement observées d'une part et simulées d'autre part. Cette différence peut être due à un biais sur les observations mêmes, à des erreurs de structure du modèle et à des erreurs dans l'estimation des variables d'entrée. Ces résultats suggèrent d'utiliser les données issues de la télédétection uniquement pour spatialiser les simulations de façon relative, en s'affranchissant de tout biais possible. Les doses préconisées par ce système seraient ensuite translatées de façon à ce que leur moyenne à la parcelle corresponde à une dose parcellaire préconisée par des méthodes plus éprouvées mais utilisant le même principe (optimisation de la marge brute sous contrainte environnementale). Enfin, il serait également judicieux de prendre en compte les contraintes techniques de la réalisation de la fertilisation azotée, qui dépendent du type d'appareil employé, dans le raisonnement spatial de la préconisation, par exemple en adaptant l'échelle de simulation à la résolution d'épandage de l'appareil.

V - Discussion et conclusion

La certitude n'est pas le fait des natures objectives.
Elles doutent facilement de leurs propres opinions et
sont prêtes à peser les arguments de l'adversaire.

Stephan Zweig, Erasme

Mais, dira-t-on, il y a des points à élucider, d'autres à résoudre.
– Je ne dis pas non ; mais en revanche il y en a beaucoup qu'il vaut mieux négliger que
rechercher : il appartient à la science d'en ignorer certains ; il y en a aussi beaucoup sur
lesquels il est plus sage de douter que de décider.
Erasme, Lettre à Dorpius

V.1 Rappel des objectifs et des principaux résultats

Cette première partie de la discussion reprend les principaux résultats des trois étapes qui ont été nécessaires au développement d'un outil de préconisation de la fertilisation azotée basé sur un modèle de culture et prenant en compte la variabilité intra-parcellaire du système sol-plante. Elle permettra de dégager des points essentiels qui seront discutés dans les parties suivantes et qui donneront lieu à des propositions de perspectives à ce travail.

La première étape consistait à évaluer la capacité du modèle de culture choisi, STICS, à préconiser la fertilisation azotée, sur un jeu de données caractérisé par des conditions pédoclimatiques variées. Ceci impliquait tout d'abord de s'assurer de la qualité de la simulation de la réponse du système sol-plante à la fertilisation azotée. Parmi les variables d'intérêt final considérées lors de la validation du modèle – le rendement, la teneur en protéines, le bilan d'azote et le reliquat d'azote à la récolte – cette dernière était la moins bien simulée, ce qui pouvait être en partie liée au fait que ses valeurs observées variaient peu. Plusieurs critères agro-environnementaux de préconisation de la fertilisation azotée ont ensuite été comparés pour différentes hypothèses de contrainte environnementale. Nous avons finalement proposé d'utiliser un critère qui choisit comme dose optimale celle qui permet de maximiser la marge brute tout en respectant un seuil concernant le bilan d'azote. L'aptitude du modèle à réaliser des préconisations selon ce critère a été évaluée par rapport à une autre méthode de préconisation classiquement utilisée (AZOBIL) et s'est révélée aussi bonne, sinon meilleure pour des contraintes environnementales plus élevées que dans la situation actuelle et non prises en compte par AZOBIL. Cet aspect sera discuté dans la partie V.2.

Dans la deuxième étape, deux méthodes de prise en compte de la variabilité intra-parcellaire ont été développées. La première a consisté en une caractérisation cartographique des propriétés des sols et en l'établissement de règles de pédotransfert permettant de générer les entrées nécessaires au modèle à partir de la carte des sols. La seconde méthode de spatialisation a mis en œuvre trois techniques d'assimilation de données issues de la télédétection adaptées à STICS et à la problématique de l'agriculture de précision. Ces trois techniques prenaient en compte une connaissance *a priori* des caractéristiques générales de la parcelle pouvant correspondre à la connaissance de l'agriculteur ou à une cartographie à faible résolution. Elles ont été comparées sur des jeux de données simulées. La méthode GLUE a été sélectionnée pour être utilisée lors de l'étape concernant les jeux de données réelles. A cette

occasion, nous avons pu montrer que les méthodes étaient efficaces en terme de correction des simulations des variables d'état mais qu'elles ne permettaient pas de retrouver les valeurs des variables d'entrée ayant servi à produire le jeu de données. Nous avons également montré que le positionnement des dates d'observation au cours du cycle de la culture jouait un rôle fondamental sur la qualité de l'assimilation, les dates tardives situées vers les mois de mai et juin ayant la plus grande efficacité. Ceci nous a amené à dire qu'il était difficile d'envisager d'utiliser les données acquises par télédétection une année *n* pour préconiser une fertilisation azotée aux mois de mars ou d'avril de l'année *n*.

Enfin, la troisième partie a permis de mettre en œuvre, comparer et valider les produits méthodologiques fournis par les deux étapes précédentes. L'évaluation de la capacité du modèle de culture renseigné par les deux approches de spatialisation a permis d'établir que la cartographie de la parcelle ne permet pas de bien restituer la variabilité intra-parcellaire, alors que l'assimilation de données issues de la télédétection donne des résultats relativement encourageants de ce point de vue, bien que les simulations des variables impliquées dans le calcul du critère de choix de la dose soient biaisées. Une discussion des méthodes de validation de la spatialisation par les deux approches sera abordée dans la partie V.3. Les résultats de l'approche par assimilation seront plus amplement commentés : la partie V.4 concerne la prise en compte des corrélations qui existent entre les variables d'entrée et les variables d'état ; la partie V.5 s'attache aux méthodes d'assimilation et à leur efficacité par rapport à un modèle de culture et la partie V.6 prend le problème de façon inverse et propose des pistes pour adapter les modèles de culture à l'assimilation de données. Au cours de la campagne 2002-2003, une expérimentation a permis d'évaluer la faisabilité de la méthode. Une carte de préconisation a été établie sur la base de simulations réalisées d'après l'approche par cartographie. Les résultats, tout en n'étant pas décisifs, indiquent un faible intérêt de la modulation. Ceci peut être dû en partie à l'approche de spatialisation utilisée, peu efficace. Une évaluation plus précise de l'intérêt de réaliser une modulation des apports d'azote sera discutée dans la partie V.7. Enfin, le point V.8 proposera des modifications de la méthode initiale permettant de prendre en compte les limites rencontrées.

V.2 Critère de choix de la fertilisation azotée

Comme le rappellent Schröder *et al.* (2000), pour être utilisable, toute stratégie de choix de la fertilisation azotée se doit d'être, à moyen terme, plus rentable économiquement que la « stratégie d'assurance ». Cette dernière consiste à mettre autant d'azote que ce qui est nécessaire à obtenir le rendement potentiel maximal sur la parcelle, c'est-à-dire celui qui peut être produit en conditions optimales de fertilisation et de climat. Tant que le prix de l'engrais restera relativement bas ou en l'absence de taxation des excès de fertilisation, cette exigence de rentabilité par rapport à la stratégie d'assurance ne sera pas aisée à atteindre.

L'utilisation d'un critère tel que ceux que nous avons proposés ne prend donc tout son sens que sous l'hypothèse de l'existence de contraintes environnementales. Supposons d'abord une contrainte sur la valeur des reliquats à la récolte. La Figure V.2-1 illustre la réponse du rendement et des reliquats à la dose d'azote décrite par les modèles statiques classiquement utilisés (*e.g.* Laurent, 2000). D'après ces modèles, le niveau des reliquats d'azote dans le sol commence à augmenter à partir du moment où la dose d'azote excède la dose de l'optimum agronomique (Chaney, 1990). Donc un critère de choix de la dose basé sur la maximisation de la marge brute suffit pour être à la valeur minimale de reliquat atteignable, car la dose de l'optimum économique est par construction inférieure à la dose de l'optimum agronomique. Ceci signifie que sous faible contrainte environnementale, un critère de décision basé sur la marge brute seule suffit. Par ailleurs, comme la valeur minimale des reliquats qui constituent un risque pour le lessivage des nitrates dépend des sols, il serait impensable de fixer un seuil qui serait partout le même. Mais le défaut principal de l'utilisation des reliquats est qu'ils ne prennent pas en compte les pertes gazeuses d'engrais notamment. Le bilan d'azote, même s'il est contestable, n'a pas ce défaut. Mais il faudrait vérifier qu'il est pertinent de fixer un seuil pour le bilan qui ne dépende pas du type de sol.

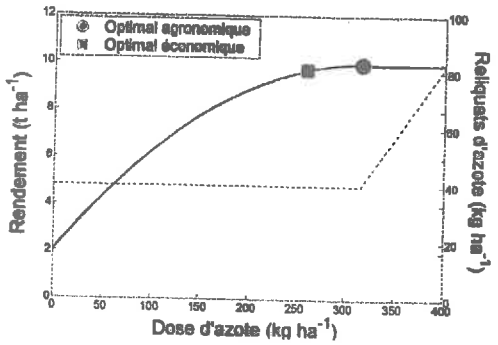


Figure V.2-1 : Courbes de réponse du rendement et des reliquats d'azote à la récolte décrite par les modèles statiques utilisés par Laurent (2000). L'optimum agronomique correspond à la dose minimale qui permet d'obtenir le rendement maximal. L'optimum économique correspond à la dose qui permet d'obtenir la marge brute maximale.

V.3 Validation de la spatialisation du modèle de culture

Les simulations de STICS faites sur la base d'une caractérisation pédologique de la variabilité intra-parcellaire associée à des règles de pédotransfert ne permettent pas de bien caractériser la variabilité spatiale observée. Nous retrouvons une des conclusions de Launay (2002), à savoir que la caractérisation fonctionnelle du sol vis à vis de la culture vue au travers des formalismes d'un modèle diverge par rapport à la caractérisation descriptive des sols (origine et nature des matériaux) complétée par des règles de pédotransfert. Une autre conclusion que l'on retrouve et soulignée par ailleurs par Batchelor *et al.* (2002) est que les propriétés concernant l'enracinement sont difficiles à estimer et peuvent en grande partie être responsables des erreurs du modèle. En effet, la simulation du fonctionnement du sol par rapport aux transferts d'eau et de nitrates est complètement dépendante de la caractérisation du sol mais aussi de l'interaction entre la croissance des racines et l'état physique du sol, très difficile à représenter.

Les simulations basées sur l'assimilation de données issues de la télédétection représentent en revanche assez bien la variabilité intra-parcellaire des variables servant à calculer le critère de choix de la dose d'azote. Mais elles sont affectées d'un biais. Plusieurs raisons à cela peuvent être avancées et seront discutées dans les parties suivantes. Il existe en outre une dispersion non négligeable qui provient pour partie de la qualité des simulations, mais il faut rappeler ici que les différences entre les échelles auxquelles les grandeurs simulées et observées sont estimées posent des problèmes méthodologiques importants. En outre, les cartes de rendement par exemple, qui servent de validation, sont elles-mêmes susceptibles de contenir des erreurs et de participer à la dispersion entre valeurs simulées et observées. Par rapport à la spatialisation, l'intérêt majeur des observations par télédétection est qu'elles couvrent de façon exhaustive l'espace, contrairement à la caractérisation du sol par les voies classiques. L'estimation des propriétés des sols sous forme de carte tend de ce fait de plus en plus à utiliser des mesures de ce type, au rang desquelles il faut également placer les méthodes développées en géophysique, comme la mesure de la résistivité du sol (Michot, 2003). Comme dans le cas de la télédétection, ces dernières sont difficiles à relier avec des propriétés classiques des sols telles que la densité apparente ou l'humidité à la capacité au champ : cela implique des calibrations ou le développement de nouveaux concepts de description des sols.

V.4 Prise en compte des corrélations

Les méthodes d'inversion de modèles et d'assimilation utilisées ici présentent de grandes similitudes. On peut dire que l'inversion concerne des modèles statiques alors que l'assimilation s'adresse à des modèles dynamiques mais qu'elles servent toutes deux à estimer les variables d'entrée d'un modèle. Aussi pourrait-on employer dans le cadre de l'assimilation des critères à minimiser développés dans celui de l'inversion et qui prennent en compte les corrélations existant entre les variables observées, entre les entrées que l'on veut estimer, voire même entre ces deux dernières (dans notre cas, on peut penser à la corrélation entre *LAI* et *QN* ou entre les humidités à la capacité au champ et au point de flétrissement). Ces critères intègrent des matrices de variance covariance entre ces différentes grandeurs, qui sont estimées par itérations successives (Viovy *et al.*, 2001). Il est à noter que la méthode GLUE permettrait de prendre en compte les corrélations entre variables observées (Makowski *et al.*, 2002). Par ailleurs, si l'information *a priori* intégrait une liaison entre variables d'entrée, elle pourrait être utilisée pour générer les simulations de départ d'après des lois liées. Dans la perspective de l'utilisation de plusieurs années de données, les corrélations entre années devraient également être prises en compte de la sorte.

Un enjeu lié au contexte de l'agriculture de précision et à la forte résolution qu'elle implique consisterait à utiliser la dépendance spatiale de certaines propriétés des sols et de la culture dans le processus d'assimilation. Dans notre étude, nous avons réalisé les assimilations de façon indépendante, pixel à pixel. Or, il serait intéressant et informatif pour estimer les valeurs caractéristiques d'un pixel donné d'utiliser l'information contenue dans ses voisins, la plupart des propriétés du sol variant généralement de façon continue et régulière. Ceci n'est cependant pas systématique mais des méthodes de détection des changements abrupts des caractéristiques spatiales sont développées (Gabriel *et al.*, sous presse).

Dans le même ordre d'idées, l'estimation des paramètres de la culture pourrait se faire de façon à ce qu'elles adoptent une valeur unique sur la parcelle, afin de respecter leur statut de caractéristique « génétique », donc non variable spatialement. Dans un tel processus, on chercherait ainsi à estimer deux types d'entrée : des paramètres et des variables au sens géographique du terme. Ceci reviendrait à réaliser le paramétrage du modèle et l'assimilation en même temps, ce qui serait particulièrement intéressant dans le cas de l'utilisation d'une variété nouvelle.

V.5 Utilisation de GLUE et d'autres méthodes d'assimilation

Au cours de ce travail, trois méthodes d'assimilation ont été utilisées, dont deux basées sur une approche stochastique. Pour traiter les données réelles, la méthode GLUE a été employée notamment car, quel que soit le nombre de pixels sur une parcelle, si l'information *a priori* est la même partout, le nombre de simulations à réaliser est le même. Seul le temps de calcul des vraisemblances augmente avec le nombre de pixels, mais, dans la situation actuelle, il reste négligeable par rapport au temps nécessaire aux simulations. Sa mise en œuvre pourrait être améliorée en utilisant les poids pour calculer les variables d'intérêt final plutôt que de les recalculer par simulation à partir des variables d'entrée tirées dans une approximation des distributions *a posteriori*. En outre, il serait également intéressant de procéder sur des données réelles aux tests qui ont été faits sur un jeu de données simulées : étude de l'influence du nombre de simulations N_G initiales réalisées, de l'erreur attribuée aux observations (σ_{LAI}^{vo} et σ_{QN}^{vo}) et de la valeur de l'information *a priori*.

Parmi les autres grandes familles de méthodes d'assimilation vues dans le chapitre introductif, des méthodes de correction de la trajectoire des variables d'état observées pourraient être utiles. L'année *n* de la préconisation, le filtre de Kalman par exemple pourrait être employé pour corriger la trajectoire de *LAI* et de *QN* grâce à des observations précoces afin de prendre en compte l'influence des variables d'entrée non permanentes, comme l'azote présent dans le sol au semis ou des conditions particulières de levée. Plus généralement, il serait intéressant de tester l'efficacité d'autres méthodes d'assimilation sur les jeux de données réelles. On peut notamment penser à des méthodes d'assimilation basées sur le calcul du gradient d'une fonction à minimiser et utilisant le modèle adjoint au modèle de culture, c'est-à-dire, schématiquement, sa dérivée par rapport à toutes les variables d'entrée.

Mais quelle que soit la méthode employée, un problème de fond demeurera. Considérons un modèle de culture composé d'une équation décrivant *LAI* et d'une autre calculant la biomasse à partir de *LAI*. Si on utilise *LAI* pour réaliser l'assimilation de données, il devient évident qu'on peut prétendre simuler correctement cette variable par la suite, mais que la qualité de la simulation de la matière sèche à partir de *LAI* va dépendre uniquement des paramètres de la seconde équation, qui ne sont pas estimables par assimilation de *LAI*, et qui sont fixés par l'information *a priori*. On a affaire à un problème « mal posé » qui explique le biais sur le rendement.

V.6 Modèles de culture adaptés à l'assimilation

La spatialisation de modèles de culture par des méthodes d'assimilation de données issues de la télédétection requiert de les adapter, car ils n'ont en général pas été conçus pour ce faire. Le premier degré d'adaptation incontournable est de permettre au modèle de simuler ce qu'on observe par télédétection. La première solution est de coupler le modèle de culture avec un modèle de transfert radiatif, ce qui a déjà été fait pour STICS (Weiss *et al.*, 2001) : c'est ce que l'on nomme une méthode ascendante. La solution que nous avons adoptée, qui est une méthode descendante, est d'utiliser le produit de l'inversion des modèles de transfert radiatif, ce qui n'affranchit pas de devoir faire simuler au modèle la teneur en chlorophylle. Dans notre approche, nous nous sommes contentés d'utiliser des relations empiriques entre *QCab* et *QN* déterminées à des dates proches ou concomitantes des dates d'observation. Mais ces mêmes relations peuvent être paramétrées en fonction du temps (Houlès *et al.*, soumis), ce qui a d'ailleurs été utilisé dans une première approche (Hollecker, 2003).

De façon plus générale, l'adéquation entre les variables utilisées pour l'assimilation et les variables d'état pourrait être améliorée en développant des modèles de culture intégrant la possibilité d'être spatialisés par des mesures de télédétection. Fondamentalement, ces dernières représentent des quantités intégrées, par exemple *QCab* plutôt que *Cab* ou *LAI* (Moulin & Guérif, soumis). Si on assimilait des données issues des mesures géophysiques, on se heurterait au même problème. Les modèles de culture adaptés à l'assimilation de ce genre de données devraient donc fonctionner plutôt à partir de variables (tant d'entrée que simulées) représentant des quantités intégrées, par exemple des réserves utiles plutôt que des humidités à la capacité au champ, ce qui éviterait du reste des difficultés liées à des compensations entre variables. Ces modèles pourraient de plus décrire les interactions spatiales sur le plan horizontal, ce qui n'est pas le cas de modèles comme STICS : on parlerait alors de modèles spatiaux et non spatialisés. En continuant ce raisonnement, on pourrait aussi proposer d'utiliser des modèles basant la production de matière sèche non pas sur l'interception de lumière par *LAI* (Monteith, 1972) mais par *QCab*. Ceci présente deux intérêts : (i) c'est cette dernière valeur qui est observable le plus directement par télédétection ; (ii) dans la plante, l'énergie lumineuse est captée non par la matière sèche des feuilles représentée par *LAI* mais bien par la quantité de chlorophylle présente et susceptible de capter la lumière.

V.7 Intérêt de la modulation

Un certain nombre d'évaluations de l'intérêt de la méthode n'a pas pu être réalisé. Compte tenu du coût potentiel de la mise en œuvre des principes de l'agriculture de précision, il est intéressant d'évaluer à partir de quelles conditions il est profitable de les appliquer.

A l'heure actuelle, les agriculteurs peuvent être intéressés par prendre en compte et caractériser la variabilité de leur parcelle sans pour autant vouloir réaliser d'apports modulés spatialement, faute du matériel nécessaire notamment. L'outil que nous avons développé pourrait également servir à préconiser une dose moyenne qui permette de satisfaire le mieux possible le critère choisi sur l'ensemble de la parcelle : ceci pose la question de savoir s'il est plus intéressant de réaliser réellement une modulation ou de se contenter de prendre en compte la variabilité pour apporter une dose constante. La réponse à cette question est naturellement dépendante de l'hétérogénéité de la parcelle (Godwin *et al.*, 2003). Elle pourrait être évaluée de la façon suivante en utilisant l'outil développé sur des données simulées. Pour plusieurs degrés d'hétérogénéité de la parcelle, on calculerait les doses optimales pixel par pixel d'une part et une dose unique permettant d'optimiser le critère à l'ensemble de la parcelle d'autre part. On étudierait ensuite le comportement de la différence entre les valeurs de critère pris pour les doses variables et pour la dose unique. On pourrait ainsi déduire le degré d'hétérogénéité de la parcelle à partir duquel il deviendrait intéressant de réaliser une modulation. Le degré d'hétérogénéité de la parcelle pourrait être évalué en terme d'hétérogénéité du rendement potentiel. De la même façon, on pourrait aussi comparer l'intérêt d'utiliser cette méthode par rapport aux méthodes plus classiques comme AZOBIL (Machet *et al.*, 1990) qui préconisent des doses constantes sans prendre en compte, dans leur version standard, la variabilité de la parcelle.

Dans le même ordre d'idée, il serait intéressant de réaliser une étude de sensibilité de la dose préconisée par la méthode aux erreurs du modèle. Ceci pourrait également se faire à partir d'un jeu de données simulées décrivant l'hétérogénéité du sol et qui servirait à établir des doses optimales. Puis, en bruitant selon des degrés divers ce jeu de données, on calculerait des doses optimales entachées d'erreur. On pourrait donc connaître la relation entre les erreurs commises sur la caractérisation de la variabilité du système sol-plante et les erreurs commises sur les doses préconisées. Ce genre d'étude pourrait servir à évaluer l'opportunité pour un agriculteur donné d'investir dans l'agriculture de précision.

V.8 Vers une version opérationnelle de la méthode

L'enchaînement de la caractérisation de la variabilité spatiale par assimilation de données issues de la télédétection et la préconisation de doses d'azote spatialement variable n'a pas été fait au cours de cette étude. Vu le nombre de simulations nécessaires à chacune de ces deux phases, cet enchaînement dans l'état actuel des choses en requerrait une quantité prohibitive. Dans la perspective d'une mise en œuvre de cette méthode, il serait donc nécessaire de réfléchir à diminuer le temps de calcul. Des voies possibles sont (i) la diminution du temps de simulation par vectorisation du modèle ou par utilisation de réseaux de neurones et (ii) l'économie du nombre de simulations à réaliser.

La méthode telle qu'initialement envisagée – caractériser la variabilité spatiale une année *n* pour réaliser des préconisations cette même année *n* – se heurte à plusieurs problèmes méthodologiques. En effet, la préconisation doit se faire vers fin mars ou début avril. Or, les acquisitions de télédétection qui permettent de caractériser l'état du couvert se font dans le domaine du visible et sont empêchées par l'ennuagement, assez fréquent à cette période de l'année en Picardie. En outre, nous avons vu que les dates qui sont les plus informatives dans le processus d'assimilation sont celles qui sont situées en milieu et en fin de cycle, soit donc après avril. Si nous rajoutons à cela l'avantage d'utiliser les données de récolte dans l'assimilation afin de mieux prédire les variables d'intérêt final, cela nous oriente vers un système d'apprentissage où la variabilité de la parcelle serait caractérisée sur une ou plusieurs années de données. Néanmoins, l'utilisation de données acquises l'année *n* permettrait de prendre en compte les variables non permanentes ou des accidents climatiques. Le nombre de dates d'observations de télédétection, sur un jeu de données simulées, a été jugé peu influent sur la qualité des assimilations. Mais ceci serait à confirmer sur des données réelles : en effet, une acquisition faite par exemple lors d'une période de stress intense pourrait conduire à exagérer la simulation des stress si elle n'était pas balancée par d'autres acquisitions. On peut également envisager d'augmenter le nombre de sources d'information utilisées dans le processus de l'assimilation, comme les données de la géophysique ou naturellement les cartes de rendement voire de teneur en protéines. Cette façon de faire est justifiée par le fait que nous avons vu que les variables d'entrée sol estimées en 1999-2000 et 2000-2001 permettent de bien simuler la variabilité observée en 2001-2002 et 2002-2003 respectivement, ce qui serait à confirmer.

V.9 Conclusion générale

La bêtise consiste à vouloir conclure.

Gustave Flaubert

Dans le contexte d'évolution actuel de l'agriculture tendant vers une recherche de pérennité des systèmes de production, des outils tels que celui développé ici pourraient trouver leur place à plusieurs niveaux. En premier lieu, ils pourraient être utilisés au niveau de l'agriculteur ou d'organismes agricoles pour prendre des décisions tactiques au sens de Morgenthaler *et al.* (2003), à savoir des décisions prises au jour le jour pour déterminer par exemple quelle dose appliquer et où. Des outils similaires mais prenant en compte plus de facteurs, également du ressort de l'agriculture de précision comme le travail du sol, l'apport de pesticides, le choix de la variété, la disponibilité des machines, la prévision des prix des intrants et des produits... pourraient servir à prendre, en amont, des décisions dites stratégiques et concernant ces facteurs. Mais il est difficilement envisageable de pouvoir prendre en compte tous les aspects des techniques culturales par des modèles, encore moins de les rassembler tous dans un seul. Pourtant, si l'on prétend utiliser des principes d'agriculture durable, ce sont toutes ces techniques qui doivent être raisonnées de façon intégrée : celles qui peuvent avoir un impact sur la pollution de l'eau ou de l'air comme la fertilisation ou l'utilisation de biocides ou celles qui permettent de limiter la dégradation des sols, comme les amendements en matière organique ou la gestion du travail du sol. Cette nécessité constitue une limite à l'utilisation de modèles, bien qu'ils restent indispensables à la prévision et à l'aide à la décision.

Ce genre d'outils qui intègrent une possibilité de faire varier le niveau des contraintes environnementales pourraient également être utilisés pour aider à la prise de décisions politiques, notamment pour évaluer l'impact de normes environnementales. Un autre intérêt serait d'évaluer l'opportunité, pour un agriculteur ou un ensemble d'agriculteurs, de s'engager dans l'agriculture de précision, en essayant de quantifier les bénéfices qu'ils pourraient en retirer. Ceci nécessiterait cependant de connaître le coût de la mise en œuvre de ces méthodes qui est difficile à réaliser compte tenu de la jeunesse de ce marché et du fait qu'il n'a pas réellement encore pris son essor. Enfin, pour reprendre la distinction établie dans l'introduction, l'intérêt de ces méthodes relativement complexes reste à démontrer par rapport aux approches indicateur, plus simples, ce qui n'est pas aisé à faire de façon expérimentale.

Si j'étais chercheur, je saurai pas quoi chercher.

JM Gouriot et al. *(2000), Brèves de Comptoir 1998-2000,*
Ed. Robert Laffont, p. 268

J'eus le grand plaisir, dans une autre chambre, de rencontrer l'agronome qui avait eu l'idée d'utiliser les cochons pour le labourage. Sa méthode, qui supprime l'emploi des charrues, des bœufs et des laboureurs, est la suivante : sur une surface à labourer d'une acre vous enfouissez à cinq ou six pouces de distance et à huit pouces de profondeur de grandes quantités de glands, châtaignes et autres fruits ou légumes dont les porcs sont particulièrement friands ; puis vous lâchez dans le champ six cents ou plus de ces animaux, lesquels, en cherchant la nourriture, retournent le sol et le laisseront à la fois préparé pour les semailles et enrichi de leurs excréments. A vrai dire, l'expérience avait montré que le procédé était coûteux et d'une pratique difficile, mais il faut croire que cette invention est susceptible encore de nombreux perfectionnements.

J Swift, Voyages de Gulliver (1726), trad. J Pons,
Ed. Gallimard, 1976, pp. 229-230

Bibliographie

AKKAL N. *Pilotage de la fertilisation azotée du blé d'hiver sur la base d'une évaluation précoce de la réflectance radiométrique et du taux de couverture du sol : application à l'agriculture de précision.* Thèse de doctorat, Institut National Agronomique Paris-Grignon, 1998, 198 p.

ANTONIADOU T. & WALLACH D. Evaluating decision rules for nitrogen fertilizer. *Biometrics*, 2000, 56, 420-426.

ARNON D.I. Coper enzymes in isolated chloroplasts ; Polyphenoloxidases in *Beta vulgaris*. *Plant Physiology*, 1949, 24, 1-15.

BARET F. *Contribution au suivi radiométrique de cultures de céréales*. Thèse de doctorat, Université d'Orsay, 1986, 182 p.

BARET F., FOURTY T. Radiometric estimates of nitrogen status in leaves. In: Lemaire G. (éd.), *Diagnosis of the Nitrogen Status in Crops*. Chapter 12. Berlin Heidelberg : Springer-Verlag ©, 1997

BARET F., WEISS M., LEROY M., ALLARD D., GARRIGUES S., BOHBOT H., BOSSENO R., DI BELLA C., ESPANA M., GOND V., GUINOT J.P., GUYON D., LELONG C., MOUGIN E., NILSON T., VINTILLAX R. VALERI: a network of sites and a methodology for the validation of land satellite products. *Remote Sensing of Environment*, soumis.

BASSO B., RITCHIE J.T., PIERCE F.J., BRAGA R.P., JONES J.W. Spatial validation of crop models for precision agriculture. *Agricultural Systems*, 2001, 68, 97-112.

BASTET G. *Estimation des propriétés de rétention en eau des sols à l'aide de fonctions de pédotransfert : développement de nouvelles approches.* Thèse de doctorat, Université d'Orléans, 1999, 215 p.

BASTIDAS L.A., GUPTA H.V., SOROOSHIAN S., SHUTTLEWORTH W.J., YANG Z.L. Sensitivity analysis of a land surface scheme using multicriteria methods. *Journal of Geophysical Research*, 1999, 19, 481-490.

BATCHELOR W.D., BASSO B., PAZ J.O. Examples of strategies to analyze spatial and temporal yield variability using crop models. *European Journal of Agronomy*, 2002, 18, 141-158.

BEAUDOIN N., NICOULLAUD B., HOULÈS V., VENET E., DEVAUX P., DELERUE J.B., GAILLARD H. Fonctions de pédotransfert associées aux bases de données sol. *Journée de restitution de l'AIP « Agriculture de Précision »*, Paris, 15 janvier 2003.

BEVEN K., BINLEY A. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 1992, 6, 279-298.

BEVEN K., FREER J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 2001, 249, 11-29.

BINDRABAN P.S. Impact of canopy nitrogen profile in wheat on growth. *Field Crop Research*, 1999, 63, 63-77.

BIRRELL S.J., SUDDUTH K.A., BORGELT S.C. Comparison of sensors and techniques for crop yield mapping. *Computers and Electronic in Agriculture*, 1996, 14, 215-233.

BLACKMORE S., MOORE M. Remedial correction of yield map data. *Precision Agriculture*, 1999, 1, 53-66.

BOOLTINK, H.W.G., van ALPHEN, B.J., BATCHELOR, W.D., PAZ, J.O., STOORVOGEL, J.J., VARGAS, R. Tools for optimizing management of spatially-variables fields. *Agricultural Systems*, 2001, 70, 445-476.

BOOIJ R., UENK D., LOKHORST C., SONNEVELD C. Monitoring crop nitrogen status in potatoes, using crop light reflection. In: *Third European Conference on Precision Agriculture*, 18-20/06/2001, Montpellier, Grenier & Blackmore eds, 893-897.

BOULAINE J. Nature et consistance des cartes pédologiques. In: *Pédologie appliquée.* Masson et C^{ie} éd., Paris, 1980, 200 p.

BOUMAN B.A.M. Linking physical remote sensing models with crop growth simulation models, applied for sugar beet. *International Journal of Remote Sensing*, 1992, 13, 2565-2581.

BRAMLEY R.G.V., COOK S.E., ADAMS M.L., CORNER R.J. Designing your own on-farm experiments: How precision agriculture can help. *Grains Research and Development Corporation, Canberra*, 1999.

BRISSON N., GARY C., JUSTES E., ROCHE R., MARY B., RIPOCHE D., ZIMMER D., SIERRA J., BERTUZZI P., BURGER P., BUSSIÈRE F., CABIDOCHE Y.M., CELLIER P., DEBAEKE P., GAUDILLIÈRE J.P., HÉNAULT C., MARAUX F., SEGUIN B., SINOQUET H. An overview of the crop model STICS. *European Journal of Agronomy*, 2003, 18, 309-332.

BRISSON N., MARY B., RIPOCHE D., JEUFFROY M.H., RUGET F., NICOULLAUD B., GATE P., DEVIENNE-BARRET F., ANTONIOLETTI R., DURR C., RICHARD G., BEAUDOIN N., RECOUS S., TAYOT X., PLENET D., CELLIER P., MACHET J.M., MEYNARD J.M., DELÉCOLLE R. 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*, 1998, 18, 311-346.

BRISSON N., RUGET F., GATE P., LORGEOU J., NICOULLAUD B., TAYOT X., PLENET D., JEUFFROY M.H., BOUTHIER A., RIPOCHE D., MARY B., JUSTES E. STICS : a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize. *Agronomie*, 2002, 22, 69-92.

BRUAND A., DUVAL O., GAILLARD H., DARTHOUT R., JAMAGNE M. Variabilité des propriétés de rétention en eau des sols : importance de la densité apparente. *Etude et gestion des Sols*, 1996, 3,1, 27-40.

BRUAND A., PÉREZ FERNÁNDEZ P., DUVAL O., QUÉTIN P., NICOULLAUD B., GAILLARD H., RAISON L., PESSAUD J.F., PRUD'HOMME L. Estimation des propriétés de rétention en eau des sols : utilisation de classes de pédotransfert après stratification texturale et texturo-structurale. *Etude et Gestion des Sols*, 2002, 9, 105-125.

CARTELAT A., GOULAS Y., LELARGE C., BARBOTTIN A., JEUFFROY M.H., CEROVIC Z.G. A new optical signature of nitrogen deficiency in wheat useful for decision support in precision agriculture. In: Programme book of the joint conference of ECPA-ECPLF. *4th European Conference on Precision Agriculture, ZALF-Agrartechnik Bornim, Allemagne, 18-20/06/2003*. Werner & Jarfe (éds), Wageningen Academic Publishers, The Netherlands, 2003, 377-379.

CASTELLI F., CONTILLO R., MICELI F. Non-destructive determination of leaf chlorophyll content in four crop species. *Journal of Agronomy & Crop Science*, 1996, 177, 275-283.

CHANEY K. The effect of nitrogen fertiliser rate on the soil nitrate content after harvesting winter wheat. *Journal of Agricultural Science (Cambridge)*, 1990, 114, 171-176.

CHAPMAN S.C., BARRETO H.J. Using a chlorophyll meter to estimate specific leaf nitrogen of tropical maize during vegetative growth. *Agronomy Journal*, 1997, 89, 557-562.

CHILÈS J.P., DELFINER P. Geostatistics: Modelling Spatial Uncertainty. In: *Probability and Statistics*, Wiley & Sons, Inter-Science, 1999.

CLAY D.E., CARLSON C.G., CHANG J., CLAY S.A., MALO D.D., ELLSBURY M.M., LEE J. Systematic evaluation of precision farming soil sampling requirements. In: Robert P.C., Rust R.H., Larson W.E. (eds), *Precision Agriculture: Proceedings of the 4th International Conference (Part A)*. ASA, CSSA, SSSA, Madison, Wisconsin. pp. 253-265.

CLEVERS J.G.P.W., LEEUWEN H.J.C. Combining directional and high spectral resolution information from optical remote sensing data for crop growth measuring. In: *Proceedings of the 6th International Symposium on Physical measurements and Signatures in Remote Sensing*, Val d'Isère (France), 17-21 janvier 1996. ISPRS-CNES, 643-650.

COMBAL B., BARET F., WEISS M., TRUBUIL A., MACÉ D., PRAGNÈRE A., MYNENI R., KNYAZIKHIN Y., WANG L. Retrieval of canopy biophysical variables from bidirectional reflectance using prior information to solve the ill-posed problem. *Remote Sensing of Environment*, 2002, 84, 1-15.

CURRAN P.J. Remote sensing of foliar chemistry. *Remote Sensing of Environment*, 1989, 30, 271-278.

DELÉCOLLE R., GUÉRIF M. Introducing spectral data into a plant process model for improving its prediction ability. In: *Proceedings of the 4th International Colloquium on Spectral Signatures of Objects in Remote Sensing*, Aussois (France), 18-22 janvier 1988.

DELÉCOLLE R., MAAS S.J., GUÉRIF M., BARET F. Remote sensing and crop production models: present trends. *ISPRS Journal of Photogrammetry and Remote Sensing*, 1992, 47, 145-161.

DEMARTY J. *Développement et application du modèle SiSPAT-RS à l'échelle de la parcelle agricole dans le cadre de l'expérience Alpilles-ReSeDA2001*. Thèse de doctorat, Université Denis Diderot, Paris VII, 2001, 217 p.

DEMARTY J., OTTLÉ C., BRAUD I., OLIOSO A., FRANGI J.P., BASTIDAS L.A., GUPTA H. Using a multi-objective sensitivity analysis approach to retrieve information on surface properties used in a SVAT model. *Journal of Hydrology*, 2004a, sous presse.

DEMARTY J., OTTLÉ C., BRAUD I., OLIOSO A., FRANGI J.P., GUPTA H., BASTIDAS L.A. Constraining a physically based SVAT model with surface water content and thermal infrared brightness temperature measurements using a multiobjective approach. *Water Resources Research*, 2004b, en révision.

DOERGE T.A. Yield map interpretation. *Journal of Production Agriculture*, 1999, 12, 54-61.

DORSAINVIL F. *Evaluation, par modélisation, de l'impact environnemental des modes de conduite des cultures intermédiaires sur les bilans d'eau et d'azote dans les systèmes de culture*. Thèse de doctorat, Institut National Agronomique Paris-Grignon, 2002, 124 p.

DUCHAUFOUR P. Systématique des sols. In: *Pédologie 1. Pédogenèse et classification.* Masson et C^ie éd., Paris, 1977, 477 p.

DUMONT K., DE BAERDEMAEKER J. In field wheat nitrogen assessment using hyperspectral imaging techniques. In: *Third European Conference on Precision Agriculture*, 18-20/06/2001, Montpellier, Grenier & Blackmore eds, 905-910.

EBELING M.E. The Dumas method for nitrogen in feeds. *Journal of the Association of Official Analytical Chemists*, 1968, 551, 766-770.

ENGEL T. Use of nitrogen simulation models for site-specific nitrogen fertilization. In: Precision agriculture '97. Volume I. Spatial variability in soil and crop. *First European Conference on Precision Agriculture, Warwick University, UK, 7-10 September 1997.* Bios Scientific Publishers Ltd, Oxford, UK, 1997, 361-369.

ESQUÍVEL M.G., FEREIRA R.B., TEIXEIRA A.R. Protein degradation in C_3 and C_4 plants subjected to nutrient starvation. Particular reference to ribulose bisphosphate carboxylase/oxygenase and glycolate oxidase. *Plant Science*, 2000, 153, 15-23.

FEIBO W., LIANGHUAN W., FUHUA X. Chlorophyll meter to predict nitrogen sidedress requirements for short-season cotton (*Gossypium hirsutum* L.). *Field Crops Research*, 1998, 56, 309-314.

FINKE P.A., WÖSTEN J.H.M., JANSEN M.J.W. Effects of uncertainty in major input variables on simulated functional soil behaviour. *Hydrological Processes*, 1996, 10, 661-669.

FOURTY T., BARET F., JACQUEMOUD S., SCHMUCK G., VERDEBOUT J. Leaf Optical Properties with Explicit Description of Its Biochemical Composition: Direct and Inverse Problems. *Remote Sensing of Environment*, 1996, 56, 104-117.

FRANCIS D.D., SCHEPERS J.S. Selective soil sampling for site-specific nutrient management. In: Precision agriculture '97. Volume I. Spatial variability in soil and crop. *First European Conference on Precision Agriculture, Warwick University, UK, 7-10 September 1997.* Bios Scientific Publishers Ltd, Oxford, UK, 1997, 119-126.

GABRIEL E., ALLARD D., BACRO J.N. Detecting zones of abrupt change: application to soil data. In: *GeoEnv IV: Geostatistics for Environmental Applications*, Kluwer academic Publisher, Dordrecht (sous presse)

GABRIELLE B., ROCHE R., ANGAS P., CANTERO-MARTINEZ C., COSENTINO L., MANTINEO M., LANGENSIEPEN M., HÉNAULT C., LAVILLEP., NICOULLAUD B., GOSSE G. *A priori* parameterisation of the CERES soil-crop models and tests against several European data sets. *Agronomie*, 2002, 22, 119-132.

GATE P. La prise de décision face à la variabilité – exemples d'applications sur les céréales à partir de la télédétection. In: *Actes du colloque UMR Cemagref-ENESAD,* Dijon (France), 29-30 mai 2000, Zwaenepoel P. (éd.), Educagri éditions, p. 383-393.

GAULTIER J.P., LEGROS J.P., BORNAND M., KING D., FAVROT J.C., HARDY R. L'organisation et la gestion des données pédologiques spatialisées : le projet DONESOL. *Revue de Géomatique*, 1993, 3, 235-253.

GHILOUFI M. *Méthodologie d'adaptation de STICS à de nouvelles cultures : application au tournesol et à la betterave.* Mémoire de Diplôme d'Agronomie Approfondie, Institut National Agronomique Paris-Grignon, 1999, 48 p.

GIJSMAN A.J., JAGTAP S.S., JONES J.W. Wading through a swamp of complete confusion: how to choose a method for estimating soil water retention parameters for crop models. *European Journal of Agronomy*, 2002, 18, 75-105.

GITELSON A.A., MERZLYAK M.N. Remote sensing estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, 1997, 18, 2691-2697.

GOMEZ O. *Participation au programme de recherche sur l'agriculture de précision de l'Unité d'Agronomie de Laon-Reims-Mons*. Mémoire d'assistant-ingénieur, Institut Supérieur d'Agriculture de Beauvais, 2002.

GONÇALVES M.C., PEREIRA L.S., LEIJ F.J. Pedo-transfer functions for estimating unsaturated hydrolic properties of Portuguese soils. *European Journal of Soil Science*, 1997, 48, 387-400.

GODWIN R.J., RICHARDS T.E., WOOD G.A., WELSH J.P., KNIGHT S.M. An economic analysis of the potential for precision farming in UK cereal production. *Biosystems Engineering*, 2003, 84, 533-545.

GRUNDY A.C., BOATMAN N.D., FROUD-WILLIAMS R.J. Effects of herbicide and nitrogen fertilizer application on grain yield and quality of wheat and barley. *Journal of Agricultural Science (Cambridge)*, 1996, 126, 379-385.

GUÉRIF M., DUKE C.L. Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data assimilation. *Agriculture, Ecosystems and Environment*, 2000, 81, 57-69.

GUÉRIF M., LAUNAY M. Local model parameters estimate using remote sensing data assimilation: application to sugar beet yield prediction and precision agriculture. In: Proceedings of the joint colloquium on *Sugar Beet and Modelling*, Lille, France, 12 Septembre 2003.

GUÉRIF M., COURAULT D., BRISSON N. Assimilation de données de télédétection dans les modèles de fonctionnement des cultures. In: *Actes de l'École-Chercheurs INRA en Bioclimatologie*, Le Croisic, 25-29 mars 1996, Lagouarde J.P. & Cruiziat P. (éds), Editions INRA, 1996, 169-191.

GUÉRIF M., BARET F., MOULIN S., BÉGUÉ A. Prise en compte de l'hétérogénéité parcellaire et de son évolution temporelle dans la gestion des interventions techniques : potentiel de la télédétection. In: *Modélisation des agro-écosystèmes et aide à la décision*, Malézieux E. (éd.), collection Repères, CIRAD, Montpellier, France, 2001a.

GUÉRIF M., BEAUDOIN N., DÜRR C., HOULÈS V., MACHET J.M., MARY B., MOULIN S., RICHARD G., BRUCHOU C., MICHOT C., NICOULLAUD N. Designing a field experiment for assessing soil and crop spatial variability and defining site-specific

management strategies. In: *Third European Conference on Precision Agriculture*, 18-20/06/2001, Montpellier, Grenier & Blackmore eds, 2001b, 677-681.

GUPTA H.V., SOROOSHIAN S., YAPO P.O. Towards improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resources Research*, 1998, 34, 751-763.

GUPTA S.C., LARSON W.E. Estimating soil water retention characteristics from particle size distribution, organic matter percent and bulk density. *Water Resources Research*, 1979, 15, 1633-1635.

HEINEMANN A.B., HOOGENBOOM G., de FARIA R.T. Determination of spatial water requirement at county and regional levels using crop models and GIS. An example for the State of Parana, Brazil. *Agricultural Water Management*, 2002, 52, 177-196.

HIKOSAKA K., TERASHIMA I. A model of the acclimation of photosynthesis in the leaves of C_3 plants to sun and shade with respect to nitrogen use. *Plant, Cell and Environment*, 1995, 18, 605-618.

HIKOSAKA K., TERASHIMA I. Nitrogen partitioning among photosynthetic components and its consequence in sun and shade plants. *Functional Ecology*, 1996, 10, 335-343.

HOEL B.O. Use of a hand-held chlorophyll meter in winter wheat: evaluation of different measuring positions on the leaves. *Acta Agriculturæ Scandinavica Sction B-Soil & Plant Science*, 1998, 48, 222-228.

HOEL B.O., SOLHAUG K.A. Effect of irradiance on chlorophyll estimation with the Minolta *SPAD*-502 leaf chlorophyll meter. *Annals of Botany*, 1998, 82, 389-392.

HOLLECKER D. *Spatialisation d'un modèle de culture pour décrire la variabilité intra-parcellaire des états du sol et de la plante*. Mémoire de fin d'études, Institut Supérieur d'Agriculture de Rhônes-Alpes, Ecole Supérieure d'Agriculture de Purpan, 2003, 57 p.

HOULÈS V., GUÉRIF M., MARY B., MACHET J.M., MAKOWSKI D. A method for optimising nitrogen fertilisation of wheat within a field based on a crop model approach. In: Programme book of the joint conference of ECPA-ECPLF. *4th European Conference on Precision Agriculture, ZALF-Agrartechnik Bornim, Allemagne, 18-20/06/2003*. Werner & Jarfe (éds), Wageningen Academic Publishers, The Netherlands, 2003, 437-438.

HOULÈS V., GUÉRIF M., MARY B., MOULIN S., MACHET J.M. Assessment of a winter wheat crop nitrogen status from measurements of leaf area index and chlorophyll content. *European Journal of Agronomy*, soumis.

HOULÈS V., MARY B., GUÉRIF M., MAKOWSKI D., JUSTES E. Evaluation of the crop model STICS to recommend nitrogen fertilisation rates according to an agro-environmental criteria. *Agronomie*, accepté.

HUET P., ALLIRAND J.M., CHELLE M., GILLOT L., JULLIEN A. Caractérisation du niveau de croissance du colza en sortie d'hiver par radiométrie visible-proche infrarouge. In: King D., Guérif M. (éds), *Prise en compte de la variabilité intra-parcellaire dans la gestion des interventions techniques*. INRA Editions, Science-Update, 2004, à paraître.

JACQUEMOUD S., BARET F. PROSPECT: a model of leaf optical properties spectra. *Remote Sensing of Environment*, 1990, 34, 75-91.

JACQUEMOUD S., BARET F., ANDRIEU B., DANSON F.M., JAGGARD K. Extraction of vegetation biophysical parameters by inversion of the PROSPECT+SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. *Remote Sensing of Environment*, 1995, 52, 163-172.

JAMAGNE M. Bases et techniques d'une cartographie des sols. *Annales Agronomiques I.N.R.A.*, 1967, hors série, vol. 18, 142p.

JAMAGNE M., BETREMIEUX R., BEGON J.-C., MORI A. Quelques données sur la variabilité dans le milieu naturel de la réserve en eau des sols. *Bulletin Technique d'Information*, 1977, 324-325, 627-641.

JAMAGNE M., HARDY R., KING D., BORNAND M. La base de données géographique des sols de France. *Etude et Gestions des Sols*, 1995, 2,3, 153-172.

JAMES I.T., GODWIN R.J. Soil, water and yield relationships in developing strategies for the precision application of nitrogen fertiliser to winter wheat barley. *Biosystems Engineering*, 2003, 84, 467-480.

JEUFFROY M.H., RECOUS S. Azodyn: a simple model simulating the date of nitrogen deficiency for decision support in wheat fertilization. *European Journal of Agronomy*, 1999, 10, 129-144.

JONES J.W., HOOGENBOOM G., PORTER C.H., BOOTE K.J., BATCHELOR W.D., HUNT L.A., WILKENS P.W., SINGH U., GIJSMAN A.J., RITCHIE J.T. The DSSAT cropping system model. *European Journal of Agronomy*, 2003, 18, 235-265.

JUSTES E., MEYNARD J.M., MARY B., PLENET D. Diagnosis using stem base extract: JUBIL method. In: Lemaire G. (éd.), *Diagnosis of the Nitrogen Status in Crops*. Chapter 10. Berlin Heidelberg : Springer-Verlag ©, 1997.

KAMPHAKE L.J., HANNAH S.A., COHEN J.M. Automated analysis for nitrate by hydrazine reduction. *Water Research*, 1967, 1, 205-216.

KARROU M., MARANVILLE J.W. Response of wheat cultivars to different soil nitrogen and moisture regimes: III. Leaf water content, conductance, and photosynthesis. *Journal of Plant Nutrition*, 1995, 18, 777-791.

KERN J.S. Evaluation of soil water retention models based on basic soil physical properties. *Soil Science Society of America Journal*, 1995, 59, 1134-1141.

KUCZERA G., PARENT E. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology*, 1998, 211, 69-85.

LAGACHERIE P., CAZEMIER D.R., MARTIN-CLOUAIRE R., WASSENAAR T. A spatial approach using imprecise soil data for modelling crop yields over vast areas. *Agriculture, Ecosystems and Environment*, 2000, 81, 5-16.

LAUNAY M. *Diagnostic et prévision de l'état des cultures à l'échelle régionale : couplage entre modèle de croissance et télédétection. Application à la betterave sucrière*. Thèse de doctorat, Institut National Agronomique Paris-Grignon, 2002, 90 p.

LAURENT F. Faut-il varier la dose d'azote en parcelles hétérogènes ? *Perspectives agricoles*, 2000, 262, 62-69.

LAUVERNET C., LE DIMET F.X., BARET F. Assimilation of remote sensing data into a canopy functioning model: the ADAM project. In: *Plant Growth Modelling and Applications*, October 2003, Beijing, 2003, pp. 305-311.

LAWLOR D.W., LEMAIRE G., GASTAL F. Nitrogen, plant growth and crop yield. In: Lea P.J., Morot-Gaudry J.F. (eds.), *Plant Nitrogen*. Berlin Heidelberg : Springer-Verlag ©, 1997.

LEGROS J.P. *Cartographies des sols. De l'analyse spatiale à la gestion des territoires.* Lausanne : Presses polytechniques et universitaires romandes, 1996, 321 p.

LEMAIRE G., GASTAL F. N uptake and distribution in plant canopies. In: Lemaire G. (éd.), *Diagnosis of the Nitrogen Status in Crops*. Chapter 1. Berlin Heidelberg : Springer-Verlag ©, 1997.

LEMAIRE G., PLENET D., GRINDLAY D. Leaf nitrogen content as an indicator of crop nutrition status. In: Lemaire G. (éd.), *Diagnosis of the Nitrogen Status in Crops*. Chapter 11. Berlin Heidelberg : Springer-Verlag ©, 1997.

LEPLASTRIER M., PITMAN A.J., GUPTA H.V., XIA Y. Exploring relationship between complexity and performance in a land surface model using the multi-criteria method. *Journal of Geophysical Research*, 2002, 107, D20, No 4443.

LE ROUX X., WALCROFT A.S., DAUDET F.A., SINOQUET H., CHAVES M.M., RODRIGUES A., OSORIO L. Photosynthetic light acclimation in peach leaves: importance of changes in mass:area ratio, nitrogen concentration, and leaf nitrogen partitioning. *Tree Physiology*, 2001, 21, 377-386.

LICHTENTHALER H.K., WELLBURN A.R. Determination of total carotenoids and chlorophyll *a* and *b* of leaf extracts in different solvents. *Biochemical Society Transactions*, 1983, 603, 591-592.

LIN H.S., McINNES K.J., WILDING L.P., HALLMARK C.T. Effects of soil morphology on hydraulic properties: II. Hydraulic pedotransfer functions. *Soil Science Society of America Journal*, 1999, 63, 955-961.

LOYCE C., RELLIER J.P., MEYNARD J.M. Management planning for winter wheat with multiple objectives (1): The BETHA system. *Agricultural Systems*, 2002, 72, 9-31.

LU C., ZHANG J. Photosynthetic CO_2 assimilation, chlorophyll fluorescence and photoinhibition as affected by nitrogen deficiency in maize plants. *Plant Science*, 2000, 151, 135-143.

MAAS S.J. Using satellite data to improve model estimates of crop yield. *Agronomy Journal*, 1988, 80, 655-662.

MACHET J.M., BEAUDOIN N., MARY B., BOFFETY D., BERNARD M. Characterization of the variability in grain production and quality within a winter wheat field. In: *Third European Conference on Precision Agriculture*, 18-20/06/2001, Montpellier, Grenier & Blackmore eds, 821-825.

MACHET J.M., DUBRULLE P., LOUIS P. AZOBIL: a computer program for fertilizer N recommendations based on a predictive balance sheet method. In: *Proceedings of the 1st ESA Congress,* Paris, 1990, p. 21-22.

MACNAB F., LAWLOR D.W., BAKER N.R., YOUNG A.T. Effects of nitrate on chloroplast composition and photosynthetic activities in wheat. In: NIJHOFF M. (éd.), *Progress in photosynthesis research, Volume 2*. Dordrecht, Netherlands, 1987.

MADSEN H. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology*, 2000, 235, 276-288.

MADSEN H., WILSON G., AMMENTORP H.C. Comparison of different automated strategies for calibration of rainfall-runoff models. *Journal of Hydrology*, 2002, 261, 48-59.

MAERTENS K., DE BAERDEMAEKER J., RAMONA H., DE KEYSER R. An Analytical Grain Flow Model for a Combine Harvester, Part II: Analysis and Application of the Model. *Journal of Agricultural Engineering Research*, 2001, 79, 187-193.

MAKOWSKI D., WALLACH D. How to improve model-based decision rules for nitrogen fertilization. *European Journal of Agronomy*, 2001, 15, 197-208.

MAKOWSKI D., JEUFFROY M.H. Estimation des paramètres des modèles de culture. In: FormaSciences INRA (éd.) *Ecole-Chercheurs, Pour une bonne Utilisation des Modèles de Culture*, Le Croisic, 14-18 octobre 2002, Paris, 2002.

MAKOWSKI D., TREMBLAY M., DEBROIZE D., LAURENT F. Epandages hétérogènes d'engrais azotés. Quel impact économique et environnemental ? *Perspectives Agricoles*, 2000, 263, 56-61.

MAKOWSKI D., WALLACH D., MEYNARD J.M. Statistical methods for predicting responses to applied nitrogen and calculating optimal nitrogen rates. *Agronomy Journal*, 2001, 93, 531-539.

MAKOWSKI D., WALLACH D., TREMBLAY M. Using a Bayesian approach to parameter estimation; comparison of the GLUE and MCMC methods. *Agronomie*, 2002, 22, 191-203.

MARY B., BEAUDOIN N., MACHET J.M. Characterisation and analysis of soil variability within two agricultural fields: the case of water and mineral profiles. In: *Third European*

Conference on Precision Agriculture, 18-20/06/2001, Montpellier, Grenier & Blackmore eds, 431-436.

MARY B., LAURENT F., BEAUDOIN N. La gestion durable de la fertilisation azotée. *Proceedings of the IIRB Colloquium, Bruxelles*, 2002, pp. 59-67.

MATSUNAKA T., WATANABE Y., MIYAWAKI T., ICHIKAWA N. Prediction of grain protein content in winter wheat through leaf color measurements using a chlorophyll meter. *Soil Science and Plant Nutrition*, 1997, 43, 127-134.

MATTHEWS R., BLACKMORE S. Using crop simulation models to determine optimum management practices in precision agriclture. In: Precision agriculture '97. Volume I. Spatial variability in soil and crop. *First European Conference on Precision Agriculture, Warwick University, UK, 7-10 September 1997.* Bios Scientific Publishers Ltd, Oxford, UK, 1997, 413-420.

MAYR T., JARVIS N.J. Pedotransfer functions to estimate soil water retention parameters for a modified Brooks-Corey type model. *Geoderma*, 1999, 91, 1-9.

McBRATNEY A.B., PRINGLE M.J. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. *Precision Agriculture*, 1999, 1, 125-152.

McKINION J.M., JENKINS J.N., AKINS D., TURNER S.B., WILLERS J.L., JALLAS E., WHISLER F.D. Analysis of a precision agriculture approach to cotton production. *Computers and Electronics in Agriculture*, 2001, 32, 213-228.

MICHOT D. *Intérêt de la géophysique de subsurface et de la télédétection multispectrale pour la cartographie des sols et le suivi de leur fonctionnement hydrique à l'échelle intraparcellaire*. Thèse de doctorat, Université Pierre et Marie Curie, Paris VI, 2003, 393 p.

MINOLTA. Specifications in detail – chlorophyll meter *SPAD* 502. *Minolta Technical Note TE102-601-01*, 1990.

MONESTIEZ P. Modèles spatialisés. In: « *Spatialisation des Modèles de Culture* », Document du séminaire INRA Département Environnement et Agronomie – INRA de Toulouse, 14-15 janvier 2002, Toulouse.

MONTEITH J.L. Solar radiation and productivity in tropical systems. *Journal of Applied Ecology*, 1972, 9, 747-766.

MORAN R. Formulae for determination of chlorophyllous pigments extracted with *N,N*-Dimethylformamide. *Plant Physiology*, 1982, 69, 1376-1381.

MORGENTHALER G.W., KHATIB N., KIM B. Incorporating a constrained optimization algorithm into remote sensing/precision agriculture methodology. *Acta Astronautica*, 2003, 53, 429-437.

MOULIN S., BONDEAU A., DELÉCOLLE R. Combining agricultural crop models and satellite observations: from field to regional scales. *International Journal of Remote Sensing*, 1998, 19, 1021-1036.

MOULIN S., GUÉRIF M. Canopy biophysical variable retrieval from hyperspectral airborne measurements: configuration of model inversion. *Remote Sensing of Environment*, soumis.

NELDER J.A., MEAD R. A simplex method for function minimization. *Computer Journal*, 1965, 7, 398-313.

NICOULLAUD B., BEAUDOIN N., ROQUE J., COUTURIER A., MAUCORPS J., KING D. Relevés pédologiques et agriculture de précision. *Journée de restitution de l'AIP « Agriculture de Précision »*, Paris, 15 janvier 2003.

NICOULLAUD B., COUTURIER A., BEAUDOIN N., MARY B., COUTADEUR C., KING D. Modélisation spatiale à l'échelle parcellaire des effets de la variabilité des sols et des pratiques culturales sur la pollution nitrique agricole. *Soumis.*

NIJBROEK R., HOOGENBOOM G., JONES J.W. Optimizing irrigation management for a spatially soybean field. *Agricultural Systems*, 2003, 76, 359-377.

OSAKI M., MORIKAWA K., MATSUMOTO M., SHINANO T., IYODA M., TADANO T. Productivity of high-yielding crops. III. Accumulation of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase and chlorophyll in relation to productivity of high-yielding crops. *Soil Science and Plant Nutrition*, 1993, 39, 399-408.

PACHEPSKY Y.A., RAWLS W.J. Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Science Society of America Journal*, 1999, 63, 1748-1757.

PAZ J.O., BATCHELOR W.D., BABCOCK B.A., COLVIN T.S., LOGSDON S.D., KASPAR T.C., KARLEN D.L. Model-based technique to determine variable rate nitrogen for corn. *Agricultural Systems*, 1999, 61, 69-75.

PENG S., GARCÍA F.V., LAZA R.C., CASSMAN K.G. Adjustment for specific leaf weight improves chlorophyll meter's estimate of rice leaf nitrogen concentration. *Agronomy Journal*, 1993, 85, 987-990.

PETTITT A.N., McBRATNEY A.B. Sampling designs for estimating spatial variance components. *Applied Statistics*, 1993, 42, 185-209.

PEYRARD N., CALONNEC A., BONNOT F., CHADŒUF J. Explorer un jeu de données sur grille de tests de permutation. *Revue de Statistiques Appliquées*, accepté.

PIEKIELEK W.P., FOX R.H. Use of a chlorophyll meter to predict sidedress nitrogen requirements for maize. *Agronomy Journal*, 1992, 84, 59-65.

PIERCE F.J., NOWAK P. Aspects of Precision Agriculture. *Advances in Agronomy*, 1999, 67, 1-85.

POCKNEE S., BOYDELL B.C., GREEN H.M., WATERS D.J., KVIEN C.K. Directed soil sampling. In: Robert P.C., Rust R.H., Larson W.E. (éds), *Precision Agriculture: Proccedings of the 3rd International Conference*. ASA, CSSA, SSSA, Madison, Wisconsin. pp. 159-168.

PRÉVOT L., CHAUKI H., TROUFLEAU D., WEISS M., BARET F., BRISSON N. Assimilating optical and radar data into the STICS crop model for wheat. *Agronomie*, 2003, 23, 297-303.

PUCKETT W.E., DANE J.H., HAJEK B.F. Physical and mineralogical data to determine soil hydraulic properties. *Soil Science Society of America Journal*, 1985, 49, 831-836.

RAJKAI K., KABOS S., van GENUCHTEN M.T. Estimation of water-retention characteristics from the buk density and particle-size distribution of Swedish soils. *Soil Science*, 1996, 161, 832-845.

RAWLS W.J., PACHEPSKY Y., SHEN M.H. Testing soil water retention estimation with the MUF pedotransfer model using data from the southern United States. *Journal of Hydrology*, 2001, 251, 177-185.

REEVES D.W. Determination of wheat nitrogen status with a hand-held chlorophyll meter: influence of management practices. *Journal of Plant Nutrition*, 1993, 16, 781-796.

ROBERT P.C. Préface. In: Actes du Colloque UMR Cemagref-ENESAD. *Agriculture de Précision, Dijon, ENESAD, France, 29-30 mai 2000.* © Educagri éditions, 2000.

RODERICK M.L., BERRY S.L., NOBLE I.R., FARQUHAR G.D. A theoretical approach to linking the composition and morphology with the function of leaves. *Functional Ecology*, 1999, 13, 683-695.

ROMANO N., SANTINI A. Effectiveness of using pedo-transfer functions to quantify the spatial variability of soil water retention characteristics. *Journal of Hydrology*, 1997, 202, 137-157.

RUGET F., BRISSON N., DELÉCOLLE R., FAIVRE R. Sensitivity analysis of a crop model, STICS, in order to choose the main parameters to be estimated. *Agronomie*, 2002, 22, 133-158.

SADLER E.J., GERWIG B.K., EVANS D.E., BUSSCHER W.J., BAUER P.J. Site-specific modelling of corn yield in the SE coastal plain. *Agricultural Systems*, 2000, 64, 189-207.

SADRAS V.O., ECHARTE L., ANDRADE F.H. Profiles of leaf senescence during reproductive growth of sunflower and maize. *Annals of Botany*, 2000, 85, 187-195.

SCHAAP M.G., FEIKE J.L., van GENUCHTEN M.T. ROSSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, 2001, 251, 163-176.

SCHAAP M.G., LEIJ F.J., van GENUCHTEN M.T. Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal*, 1998, 62, 847-855.

SCHEINOST A.C., SINOWSKI W., AUERSWALD K. Regionalization of soil water retention curves in a highly variable soilscape, I. Developing a new pedotransfer function.. *Geoderma*, 1997, 78, 129-143.

SCHRÖDER J.J., NEETESON J.J., OENEMA O., STRUIK P.C. Does the crop or the soil indicate how to save nitrogen in maize production? Reviewing the state of the art. *Field Crops Research*, 2000, 66, 151-164.

SEELAN S.K., LAGUETTE S., CASADY G.M., SEIELSTAD G.A. Remote sensing applications for precision agriculture: A learning community approach. *Remote Sensing of Environment*, 2003, 88, 157-169.

SHADCHINA T.M., DMITRIEVA V.V., MORGOUN V.V. Interrelation between nitrogen status of the plants, leaf chlorophyll and grain yield in various winter wheat cultivars. *Acta Agronomica Hungarica*, 1998, 46, 25-34.

SHAFFER M.J., BRODAHL M.K. Rule-based management for simulations in agricultural decision support systems. *Computers and Electronics in Agriculture*, 1998, 21, 135-152.

SHANGGUAN Z., SHAO M., DYCKMANS J. Effects of nitrogen nutrition and water deficit on net photosynthetic rate and chlorophyll fluorescence in winter wheat. *Journal of Plant Physiology*, 2000, 156, 46-51.

SHIRAIWA T., SINCLAIR T.R. Distribution of nitrogen among leaves in soybean canopies. *Crop Science*, 1993, 33, 804-808.

SINOWSKI W., SCHEINOST A.C., AUERSWALD K. Regionalization of soil water retention curves in a highly variable soilscape, II. Comparison of regionalization procedures using a pedotransfer function. *Geoderma*, 1997, 78, 145-159.

SIVASANKAR A., LAKKINENI K.C., VANITA JAIN, KUMAR P.A., ABROL Y.P. Differential response of two wheat genotypes to nitrogen supply. I. Ontogenic changes in laminae growth and photosynthesis. *Journal of Agronomy & Crop Sciences*, 1998, 181, 21-27.

SOROOSHIAN S., GUPTA V.K. Model calibration. In: Singh V.P. (éd.), *Computer Models of Watershed Hydrology*. Water Resources Publication, 1130 p.

STAFFORD J.V., AMBLER B., BOLAM H.C. Cut width sensors to improve the accuracy of yield mapping systems. In: Precision agriculture '97. Volume II. Technology, IT and Management. *First European Conference on Precision Agriculture, Warwick University, UK, 7-10 September 1997*. Bios Scientific Publishers Ltd, Oxford, UK, 1997, 119-126.

STANFORD G. Rationale for optimum nitrogen fertilization in corn production. *Journal of Environmental Quality*, 1973, 2, 159-166.

TARDIEU F., MANICHON H. Caractérisation en tant que capteur d'eau de l'enracinement du maïs en parcelle cultivée. II. - Une méthode d'étude de la répartition verticale et horizontale des racines. 1986, 6, 415-425.

TIETJE O., TAPKENHINRICHS M. Evaluation of pedo-transfer functions. *Soil Science Society of America Journal*, 1993, 57, 1088-1095.

UŽÍK M., ŽOFAJOVÁ A. Chlorophyll and nitrogen content in leaves of winter wheat at different genotypes and fertilization. *Rostlinná Výroba*, 2000, 6, 237-244.

VAN ALPHEN B.J. A case study on precision nitrogen management in Dutch arable farming. *Nutrient Cycling in Agroecosystems*, 2002, 62, 151-161.

VAUCLIN M., VIEIRA S.R., VACHAUD G., NIELSEN D.R. The use of cokriging with limited soil observations. *Soil Science Society of America Journal*, 1983, 47, 175-184.

VERHOEF W. Light scattering by leaf layers with application to canopy reflectance modelling: The SAIL model, *Remote Sensing of Environment*, 1984, 16, 125-141.

VIDAL D., BERGARECHE M.C. Protein and chlorophyll content in leaves of *Triticum aestivum* L. cv. Kolibri and cv. Astral. In: Metzner H. (éd.), *Photosynthesis and plant productivity. Joint meeting of the O.E.C.D. and Studiezentrum Weikersheim, Ettligen Castle (Allemagne), 11-14 octobre*. Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart, RFA, 1983, 63-67.

VIDAL I., LONGERI L., HETIER J.M. Nitrogen uptake and chlorophyll measurements in Spring Wheat. *Nutrient Cycling in Agroecosystems*, 1999, 55, 1-6.

VIOVY N., FRANÇOIS C., BONDEAU A., KRINNER G., POLCHER J., KERGOAT L., DEDIEU G., DE NOBLET N., CIAIS P., FRIEDLINGSTEIN P. Assimilation of remote sensing measurements into the ORCHIDEE/STOMATE DGVM biosphere model. In: *Eight International Congress on Physical Measurements in Remote Sensing*, 8-12/01/2001, Aussois, France, CNES éd., 713-718.

VOUILLOT M.O., HUET P., BOISSARD P. Early detection of N deficiency in wheat crop using physiological and radiometric methods. *Agronomie*, 18, 117-130.

VRINDTS E., REYNIERS M., DARIUS P., DE BAERDEMAEKER J., GILOT M., SADAOUI Y., FRANKINET M., HANQUET B., DESTAIN M.-F. Analysis of soil and crop properties for precision agriculture for winter wheat. *Biosystems Engineering*, 2003, 85, 141-152.

WALLACH D. Section 2 - Evaluation d'un modèle de culture. In: *« Pour une bonne utilisation des modèles de culture »*, Document de l'Ecole-Chercheurs INRA Département Environnement et Agronomie - FormaSciences, 14-18 octobre 2002, Le Croisic.

WALLACH D., GOFFINET B., BERGEZ J.E., DEBAEKE P., LEENHARDT D., AUBERTOT J.N. Parameter estimation for crop models: a new approach and application to a corn model. *Agronomy Journal*, 2001, 93, 757-766.

WEISS M., TROUFLEAU D., BARET F., CHAUKI H., PRÉVOT L., OLIOSO A., BRUGUIER N., BRISSON N. Coupling canopy functioning and radiative transfer models for remote sensing data assimilation. *Agricultural and Forest Meteorology*, 2001, 108, 113-128.

WELSH J.P., WOOD G.A., GODWIN R.J., TAYLOR J.C., EARL R., BLACKMORE S., KNIGHT S.M. Developing strategies for spatially variable nitrogen application in cereals, Part II: Wheat. *Biosystems Engineering*, 2003, 84, 495-511.

WENDROTH O., REUTER H.I., KERSEBAUM K.C. Predicting yield of barley across a landscape: a state-space modelling approach. *Journal of Hydrology*, 2003, 272, 250-263.

WILLMOTT C.J. On the validation of models. *Physical Geography*, 1981, 2, 184-194.

WÖSTEN J.H.M., PACHEPSKY Ya.A., RAWLS W.J. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 2001, 251, 123-150.

WÖSTEN J.H.M., SCHUREN C.H.J.E., BOUMA J., STEIN A. Functional sensitivity analysis of four methods to generate soil hydraulic functions. *Soil Science Society of America Journal*, 1990, 54, 832-836.

YAPO P.O., GUPTA H.V., SOROOSHIAN S. Multi-objective global optimisation for hydrologic models. *Journal of Hydrology*, 1998, 204, 83-97.

YODER B., PETTIGREW-COSBY R.E. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. *Remote Sensing of Environment*, 1995, 53, 199-211.

ZHANG N., WANG M., WANG N. Precision agriculture – a worldwide overview. *Computers and Electronics in Agriculture*, 2002, 36, 113-132.

ZWAENEPOEL P., LE BARS J.M. L'agriculture de précision. *Ingénieries – EAT*, 1997, 12, 67-79.

Abstract

The limitation of nitrogen fertilisation impacts on air and soil quality, as long as farmers' income optimisation, implies to take into account inner-field soil variability and to recommend spatially variable nitrogen rates. This work, belonging to an INRA project on precision agriculture, put this aim into practice for winter wheat. A tool based on the crop model STICS was developed. Two fields located in Picardy and four years measurements formed the experimental base of this study.

During a first step, we established a method to recommend nitrogen rates thanks to the model and according to several aims formalised into an "agro-environmental criterion". The crop model ability to describe the response of the soil-crop system to the nitrogen rate was first evaluated for several soils and climatic conditions. Then several agro-environmental criteria were compared. To combine agronomic, economic and crop quality concerns, all the criteria were based on the calculation of the gross margin. An environmental constraint was added following several hypothesis because no law exists at the moment. They combined (i) the two possible environmental variables: the mineral nitrogen in soil at harvest or the nitrogen balance and (ii) two solutions for taking into account the constraint: using a tax or a threshold. The interest of these four different criteria was tested on a database of nitrogen trials. The model ability to make nitrogen fertilisation recommendations according to them was evaluated. Climatic scenarios enabled to cope with climate uncertainty.

In the second step, two different methods to spatialise the crop model simulations at the inner-field scale were implemented. The first one consisted in drawing a soil map and to build pedotranfer functions. The goal of pedotransfer functions was here to link the map to the model inputs. The second one used data derived from remote sensing (leaf area index and nitrogen absorbed by the crop) to estimate soil inputs of the model thanks to data assimilation methods. Three assimilation methods were compared on a simulated dataset. They all used prior information on model inputs, which allows to estimate more inputs and avoid that they take abnormal values. One of them was selected to be used on the real dataset.

The third step combined the first two on the experimental dataset. The two spatialisation approaches were compared. The data assimilation method was the best to describe inner-field variability, even if yield simulations were biased. A nitrogen fertilisation map was established in 2003 on the basis of the soil map approach. The experiment led to a first evaluation of the opportunity to spatially recommend nitrogen fertilisation.

Key words: precision agriculture, crop model, winter wheat, recommandation of nitrogen fertilisation, agro-environmental criterion, spatialisation, soil mapping, remote sensing data assimilation.

Résumé en Français au dos.

Bibliothèque
I.N.R.A. - C.S.E.
Unité Climat Sol et Environnement
Bâtiment Climat
Site Agroparc
F. 84914 AVIGNON CEDEX 9
FRANCE

Résumé

Afin de limiter les conséquences de la fertilisation azotée sur la qualité de l'air et de l'eau et d'optimiser le revenu agricole, l'hétérogénéité des parcelles doit être prise en compte en préconisant des doses d'azote variables à l'échelle intra-parcellaire. Ce travail, initié au sein d'un projet de l'INRA concernant l'agriculture de précision, a mis en œuvre ce principe par rapport à la fertilisation du blé d'hiver. Nous avons pour cela développé un outil basé sur l'utilisation du modèle de culture STICS. Le support expérimental principal de l'étude était constitué de deux parcelles agricoles dans le département de l'Aisne suivies sur quatre années.

Une première étape consistait à établir une méthode de préconisation des doses d'azote grâce au modèle selon des objectifs formalisés par un critère agro-environnemental. Pour ce faire, la capacité du modèle à décrire la réponse du système sol-plante à la dose d'azote a d'abord été évaluée sur une large gamme de conditions pédoclimatiques. Puis plusieurs critères ont été proposés et comparés. Afin de combiner des impératifs agronomiques, économiques et de qualité de la récolte, tous les critères étaient basés sur le calcul de la marge brute. Une contrainte d'ordre environnemental a été intégrée selon plusieurs formalismes hypothétiques, en l'absence de législation actuelle. Ils combinaient l'utilisation des reliquats ou du bilan d'azote comme variable environnementale et deux formes d'application de la contrainte, par une taxation ou par l'imposition d'un seuil à ne pas dépasser. L'intérêt de ces différents critères a été comparé sur un jeu de données d'essais de doses d'azote. L'aptitude du modèle à réaliser des préconisations pertinentes avec ces critères a ensuite été testée. L'incertitude climatique a été prise en compte en utilisant des scénarios de climats possibles pour réaliser les simulations.

La deuxième étape a permis de mettre en place deux grands types d'approche permettant de spatialiser les simulations du modèle à l'échelle intra-parcellaire. La première d'entre elle a consisté à dresser un inventaire cartographique des sols associé à des règles de pédotransfert qui ont permis d'établir le lien avec les données d'entrée nécessaires au modèle. La seconde approche de spatialisation a utilisé des mesures issues de la télédétection pour estimer certaines variables d'entrée du modèle par assimilation de données (indice de surface foliaire et quantité d'azote absorbée par la culture). Lors de cette étape, trois techniques ont été comparées sur des jeux de données simulées. Elles permettaient toutes d'intégrer une information *a priori* sur les variables d'entrée, ce qui autorise à en estimer davantage et évite qu'elles ne prennent des valeurs aberrantes. Une d'entre elles a été retenue pour être utilisée sur le jeu de données réelles.

La troisième étape a mis en œuvre les deux premières sur les parcelles expérimentales. Les deux approches de spatialisation ont été comparées. L'approche utilisant l'assimilation de données s'est révélée être la plus apte à décrire la structure spatiale de la variabilité intra-parcellaire, bien que les simulations de rendement qu'elle permet soient affectées d'un biais. Une carte de préconisation de la fertilisation azotée a été établie pour la saison 2003 sur la base de l'approche cartographique et le dispositif expérimental a permis d'évaluer en première approche l'intérêt de réaliser une fertilisation modulée.

Mots clefs : agriculture de précision, modèle de culture, blé d'hiver, préconisation de la fertilisation azotée, critère agro-environnemental, spatialisation, cartographie des sols, assimilation de données de télédétection.