# Gestion de séries pluriannuelles de cartes de rendement

B. Panneton, M. Brouillard et T. Piekutowski

*Centre de recherche et de développement en horticulture, Agriculture et Agroalimentaire Canada, St-Jean-sur-Richelieu, Québec, Canada J3B 3E6.*

Panneton, B., Brouillard, M. et Piekutowski, T. 2002. **Gestion de séries pluriannuelles de cartes de rendement**. Canadian Biosystems Engineering/Le génie des biosystèmes au Canada **44**: 7.23-7.29. En agriculture de précision, on cherche à distinguer la variabilité dans les rendements qui est associée aux conditions particulières du terrain de celle qui résulte de pratiques culturales ou de conditions climatiques annuelles. Une méthode permettant de faire la synthèse de données de rendement s'étalant sur plusieurs années en cherchant à isoler l'effet des conditions locales de terrain a été développée. Il s'agit d'un algorithme de classification supervisée faisant appel à la logique floue. Les données annuelles de rendement sont traitées à l'aide de fonctions d'appartenance qui peuvent être appliquées de sorte que les facteurs liés à l'historique de la parcelle ou aux pratiques culturales peuvent être pris en compte. Pour chaque année, il en résulte un jeu de trois cartes d'appartenance. Chacune de ces trois cartes est combinée aux cartes correspondantes d'autres années en utilisant des règles plus ou moins élaborées. Finalement, des règles de classification sont appliquées pour produire une carte-synthèse montrant les régions du champ ou le rendement est stable dans le temps (SDT) et faible, SDT et typique pour la parcelle, SDT et élevé et instable dans le temps (IDT). **Mots clés**: carte, rendement, agriculture de précision, logique floue, classification.

In precision farming, it is useful to separate yield variations associated with local terrain conditions to those resulting from cultural practices or annual climatic conditions. A method for combining yield maps from several years that isolates the effect of local terrain conditions has been developed based on fuzzy logic concepts. Annual yield data are transformed using a set of membership functions that can take into account the effect of cultural practices or yield history factors. For each year, three membership maps result. Each of these maps is combined with corresponding maps built from yield data for other years using more or less elaborated rules. Finally, classification rules are applied to segment the field into four categories: stable in time and high yield (SIT-high), SIT-average, SIT-low and unstable in time. **Keywords**: map, yield, precision farming, fuzzy logic, classification.

## INTRODUCTION

Le rendement représente l'extrant des productions végétales et cette mesure est donc essentielle pour évaluer la pertinence des pratiques culturales. Avant l'éclosion de l'agriculture de précision, le rendement était évalué à l'échelle de la parcelle. L'agriculture de précision qui, en anglais, a déjà reçue l'appellation "farming-by-the-foot" a considérablement réduit l'échelle à laquelle le rendement peut être évalué. Cette question d'échelle n'est pas simple et ce n'est pas le capteur de rendement qui peut enregistrer une donnée à chaque mètre le long du parcours de la machine qui doit fixer l'échelle pertinente d'analyse pour les données de rendement. Pour une parcelle, l'identification des zones à l'intérieur desquelles la culture répond de façon homogène est un moyen permettant de mieux définir l'échelle pertinente (Lark et Stafford 1998).

Un certain nombre d'approches ont été proposées pour faire une segmentation de la parcelle à partir des données de rendement. L'inspection visuelle d'une carte annuelle de rendement permet d'identifier des zones où le rendement est soit élevé ou plus faible. Cette répartition résulte de l'effet de facteurs intrinsèques (e.g. classe de sol) et de facteurs annuels (e.g. précipitations) (Nolin et al. 1997). Il n'est donc pas possible d'isoler l'effet des facteurs intrinsèques en ne se basant que sur une seule année de données. Blackmore et al. (1999) proposent de faire la synthèse de données de rendement s'étalant sur plusieurs années en calculant la moyenne dans le temps pour chaque élément d'une grille superposée à la parcelle et la subdivisant en petites unités (de l'ordre de 10 m x 10 m). Les données de rendement sont préalablement mises à l'échelle en divisant le rendement pour chaque élément de la grille par le rendement moyen de la parcelle. Le coefficient de variation correspondant à chacune des valeurs moyennes calculées est aussi enregistré. Il en résulte deux cartes (moyenne et coefficient de variation) à partir desquelles des zones à rendement SDT-faible, SDT-élevé et IDT sont identifiées. Broos et al. (1999) proposent une approche similaire mais où les données brutes de rendement sont d'abord transposées de façon linéaire sur une échelle de 1 à 5 qui couvre la gamme des valeurs de rendement pour l'année. Ces valeurs annuelles sont alors additionnées par chacun des éléments de la grille. Finalement, une classification est obtenue en posant que les sommes de faible valeur correspondent à des points où le rendement est SDT-faible, les valeurs élevées à un rendement SDT-élevé et les valeurs intermédiaires à un rendement IDT. Lark et Stafford (1996) ont développé une méthode de classification non-supervisée faisant appel aux ensembles flous. Le résultat est une carte où chaque élément de la grille est assigné à une classe. La méthode de traitement comprend un algorithme et des règles de décision pour établir le nombre de classes. Cette méthode ne demande pas de poser des hypothèses a priori. Elle cherche à identifier les zones (contiguës ou pas) dans la parcelle où la réponse de la culture est homogène. La méthode ne permet pas de connaître ce qui induit la réponse ni la nature de cette réponse. Une analyse subséquente doit être initiée pour le découvrir.

Cet article décrit une nouvelle méthode de classification de données de rendement permettant de faire la synthèse de séries pluriannuelles de cartes. La méthode est décrite en détail à

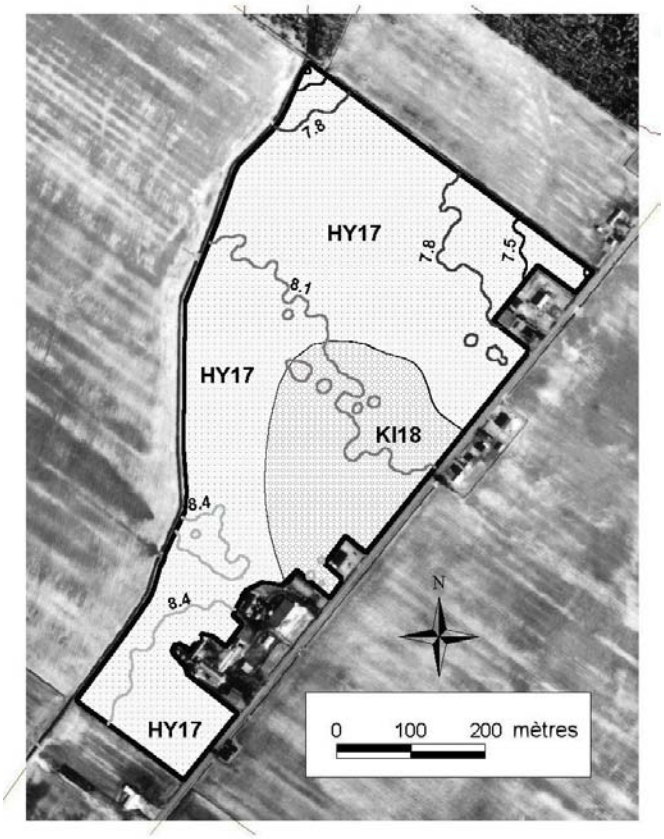


**Fig. 1. Site expérimental. Les lignes sinueuses sont des lignes de niveau identifiées par la hauteur en mètre. La trame montre les zones pédologiques (HY17 et KI18).**

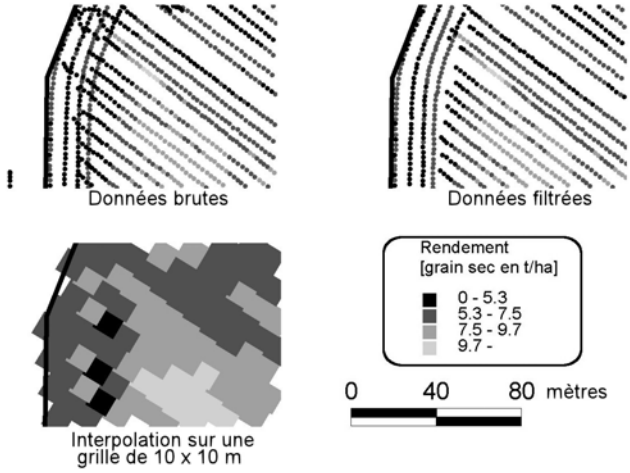


**Fig 2. Illustration du processus de traitement des données brutes pour obtenir des données sur un grille régulière.**

l'aide d'un jeu de données réel s'étalant sur trois années. L'objectif de la méthode de classification est de segmenter la parcelle en zones où le rendement est stable et instable dans le temps. Pour les zones où le rendement est stable la segmentation permet d'identifier des zones où le rendement est bas, typique ou élevé pour la parcelle.

## DONNÉES ANNUELLES DE RENDEMENT

Les données de rendement servant à illustrer la méthode proviennent d'un champ située dans la plaine du Saint-Laurent au sud-est de Montréal à La Présentation, Québec. Au nord-ouest la bordure du champ suit le tracé d'un ruisseau (Fig. 1). Le terrain a une légère pente du sud-ouest vers l'aval au nord-est. La pente moyenne est d'environ 0.15%. Dans la région, les sols sont principalement composés de gleysols humiques orthiques provenant de dépôts lacustres. La série de Kierkoski (KI18) est une argile limoneuse alors que la série St-Hyacinthe constituant la plus grande partie de la parcelle (HY17) est un loam limoneux argileux. Typiquement, ces sols sont dépourvus de fragments grossiers et mal drainés avec une structure de surface déficiente. Cultivés convenablement, ces sols donnent des rendements moyennement élevés à élevés.

Les données de rendement portent sur trois années consécutives (1996-1998) de culture de maïs-grain. Elles ont été acquises à l'aide du système GreenStar™ (John Deere, Moline, IL). Les données ont été exportées en format ASCII à l'aide du logiciel JDMap™ version 2.1 (John Deere, Moline, IL). Le traitement subséquent des données a été réalisé à l'aide du logiciel Arcview 3.2™ (ESRI Canada, Toronto, ON). Les données brutes de rendement au sortir du champ doivent être filtrées pour éliminer des mesures erronées ou douteuses (Blackmore et Moore 1999). Il faut d'abord introduire un décalage dans les données pour tenir compte du temps de transit des produits récoltés entre la barre de coupe (*i.e.* leur position sur le terrain) et le capteur de rendement installé sur la machine de récolte. Cette étape a été réalisée en premier. Lorsque le décalage est approprié, la majorité des points associés aux virages disparaissent (Fig. 2). C'est en maximisant l'élimination des points correspondant aux virages qu'il a été possible de sélectionner le délai approprié. Le rendement humide a été ensuite converti en rendement sec (14.5% d'humidité pour le maïs-grain). Finalement, une série de filtres a été utilisée pour éliminer des données:

- Données dont la localisation est à l'extérieur de la parcelle;
- Données pour lesquelles il n'y a pas de correction différentielle du GPS;
- Données erronées sur la largeur de coupe (normalement, l'opérateur doit entrer cette donnée lors de la récolte. Cette étape est souvent négligée.);
- Données dont la position est trop rapprochée de la donnée précédente (moins de 250 mm). Dans ces cas, le rendement calculé (poids mesuré ÷ (largeur×distance)) est imprécis;
- Données pour lesquelles l'humidité mesurée du grain est trop haute (>40%) ou trop basse (<14.5%);
- Données pour lesquelles le rendement est nul ou trop élevé (maïs: >20 t/ha);
- Une moyenne et un écart-type (E.T.) ont été calculés en utilisant les points restants. Les points pour lesquels le rendement etait en dehors de la plage comprise entre ± 3 E.T. autour de la moyenne ont été éliminés.

Par la suite, les données restantes ont été utilisées pour calculer un semi-variogramme (Webster et al. 1990). Un modèle

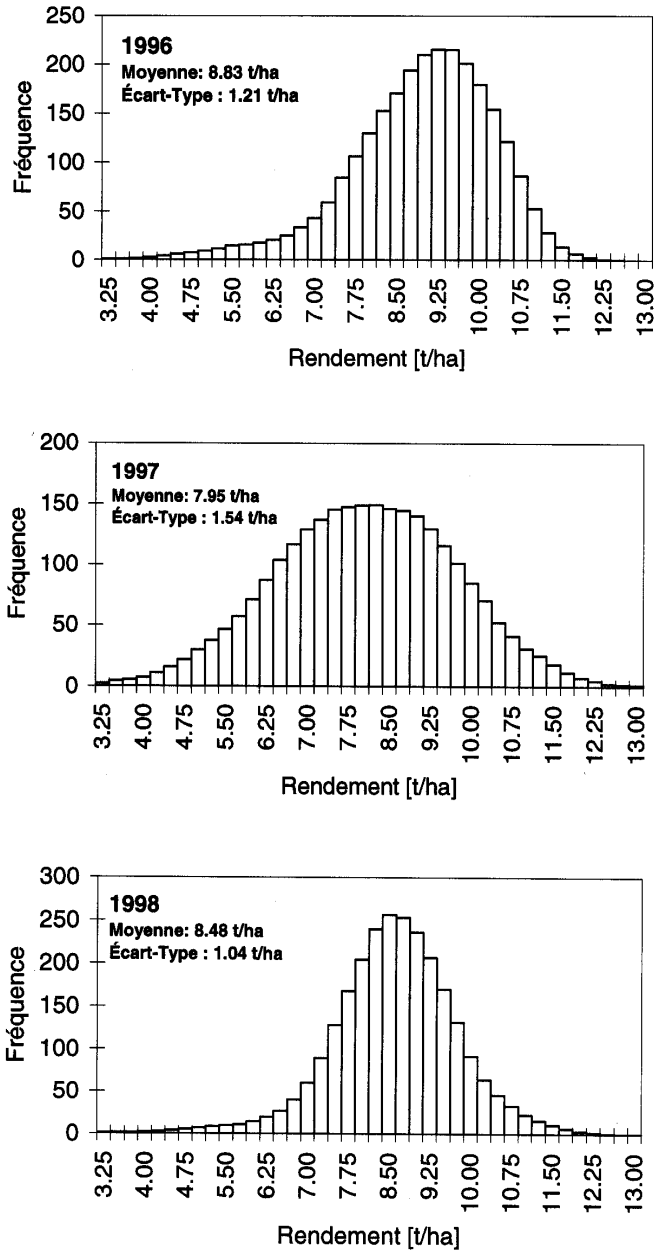


**Fig. 3. Histogrammes de rendement filtrés par une moyenne mobile.**

a été ajusté au semi-variogramme. Ce modèle a servi à interpoler par krigeage (Webster et al. 1990) les données de rendement afin de les transposer sur une grille régulière dont chacun des éléments mesure 10 m x 10 m (Fig. 2). Pour les données utilisées dans cet article, un modèle exponentiel a été utilisé.

Les histogrammes de rendement sont présentés à la Fig. 3. Pour un rendement donné (axe horizontal), la hauteur de l'histogramme (fréquence) est proportionnelle au nombre total de points où le rendement mesuré est compris entre les bornes définissant la largeur des colonnes sur l'axe horizontal. Le rendement moyen était de 8.83, 7.95 et 8.48 t/ha pour les années 1996, 1997 et 1998 respectivement (Fig. 3) et les écart-types étaient respectivement 1.21, 1.54 et 1.04 t/ha. Pour les the trois années, le coefficient d'asymétrie était négatif (-0.54, -0.07 et –1.00 pour 1996, 1997 et 1998 respectivement) indiquant que les surfaces où le rendement était inférieur au mode de la distribution étaient plus grandes que les surfaces où le rendement était supérieur au mode.

## CLASSIFICATION DES DONNÉES DE RENDEMENT

Le traitement de séries pluri-annuelles de cartes de rendement a été abordé par le biais de méthodes de prise de décision. Une décision est le résultat d'un processus par lequel on arrive à choisir parmi des alternatives. Dans l'interprétation des cartes de rendement, l'identification de zones ayant des caractéristiques stables à mesure que les saisons passent est importante. De plus, pour ces zones stables, il est intéressant de déterminer si le rendement est généralement élevé ou faible. À partir de ces considérations, quatre classes ont été définies:

- Rendement stable dans le temps et faible (SDT - faible);
- Rendement SDT - typique (typique pour la parcelle);
- Rendement SDT - élevé;
- Rendement instable dans le temps (IDT).

Trois classes ont été réservées pour les zones où le rendement tend à être stable année après année. Par rapport à une classification à deux niveaux (sous et au-dessus de la moyenne), le but est de faire ressortir plus clairement les zones où le rendement est nettement limité pour simplifier le processus de diagnostic. Ces quatre classes constituent l'ensemble des alternatives. Le processus de décision doit donc permettre de classer chaque point du champ dans une de ces quatre classes.

Le processus de prise de décision comprend trois étapes: classification des données sur une base annuelle en rendement bas, typique et élevé; agglomération des données de plusieurs années point par point et classification finale.

Pour effectuer la classification sur une base annuelle, il faut d'abord définir les notions de rendement bas, typique et élevé. Ces notions peuvent être définies de plusieurs façons. Pour éviter une définition rigide, l'approche par logique floue (Zhang et al. 2000) a été retenue. Avec cette approche, chaque point d'une carte annuelle de rendement reçoit trois valeurs indiquant son degré d'appartenance à une des trois catégories.

Le degré d'appartenance a été calculé à l'aide de deux fonctions trapézoïdales et d'une fonction triangulaire (Fig. 4). Ces fonctions ont été définies l'aide de trois repères obtenus à partir de l'histogramme des rendements pour la parcelle considérée (a, b et p sur la Fig. 4). Le mode de l'histogramme, (p sur la Fig. 4), est la valeur du rendement la plus fréquente dans la parcelle. Cette valeur est immédiatement sous le pic principal de l'histogramme. Habituellement, les histogrammes de rendement ont un pic majeur qui est bien défini. Des pics secondaires peuvent apparaître. Lorsqu'un pic secondaire est rapproché du pic principal et que leurs amplitudes sont comparables, les données de l'histogramme sont lissées à l'aide d'un filtre passe-bas (moyenne mobile) ayant une capacité filtrante suffisante pour qu'un seul pic principal subsiste dans l'histogramme résultant. L'utilisation du mode de l'histogramme plutôt que de la valeur moyenne du rendement est indifférente lorsque l'histogramme est parfaitement symétrique. En pratique,

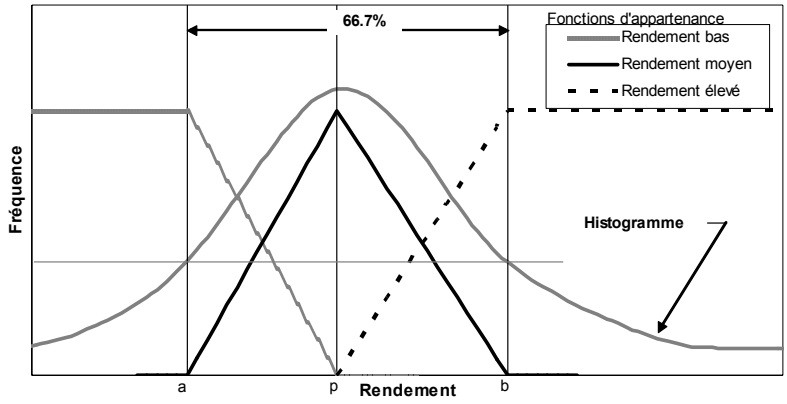


**Fig. 4. Illustration des fonctions d'appartenance et de leurs paramètres. L'histogramme est hypothétique.**

cela est rare. Généralement, la forme générale de l'histogramme comprend deux composantes: une distribution d'allure symétrique centrée autour du mode à laquelle s'ajoute une longue queue de faible amplitude s'étendant vers des valeurs de faible rendement. L'impact de cette asymétrie est que le rendement moyen est inférieur au mode. Dans ce contexte, le mode est un repère plus approprié pour définir le rendement typique du champ, c'est à dire le rendement que l'on rencontre le plus souvent à l'intérieur de la parcelle.

Les paramètres, a et b ont été obtenus à partir de la segmentation de l'histogramme (Fig. 4). Cette segmentation est analogue à celle qui est faite sur des distributions symétriques en utilisant une mesure de dispersion (habituellement l'écart-type) pour définir un intervalle autour de la moyenne à l'intérieur duquel une proportion définie de la population se retrouve. Par exemple, pour une distribution normale, on retrouve 68% de la population à ±1 écart-type autour de la moyenne. Pour une distribution qui n'est pas symétrique, on obtient une définition analogue en identifiant une fréquence de référence pour laquelle, la surface sous la courbe de l'histogramme, dans sa portion où les fréquences sont

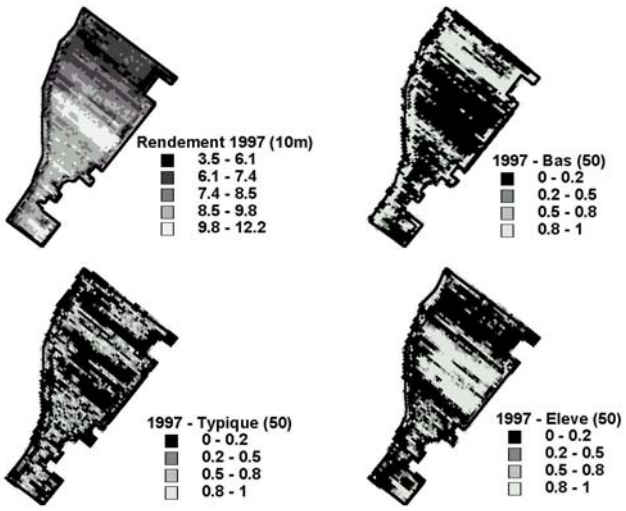


**Fig. 5. Les trois cartes d'appartenance pour 1997. Surface de référence = 50%.**

**Tableau 1. Équivalence des opérations de base - logique binaire versus logique floue.**

| Logique booléenne (ou binaire) | Logique floue |
|---|---|
| a ET b | Min (a, b) |
| a OU b | Max (a, b) |
| | Moyenne (a, b) |

supérieures à la fréquence de référence, est égale à une proportion définie. Sur la Fig. 4 par exemple, les lignes verticales aux valeurs a et b font intersection avec la courbe de l'histogramme, de part et d'autre du mode, à la même valeur de la fréquence. La portion de l'histogramme comprise entre ces deux lignes regroupent 2/3 des points de rendement. Cette portion de la surface sous la courbe (2/3) est au choix de l'utilisateur. Une fois qu'elle est choisie, la définition exacte des trois fonctions d'appartenance peut être calculée de façon automatique pour chacune des années de rendement. Cette portion de la surface sous la courbe de l'histogramme est donc un paramètre de la méthode de classification. Dans la suite du texte, il est appelé la surface de référence.

En appliquant les fonctions d'appartenance pour chaque carte annuelle de rendement, trois cartes ont été produites, une pour chaque fonction d'appartenance. Sur ces cartes (exemple à la Fig. 5), les points ont une valeur qui varie entre 0 et 1. Plus la valeur se rapproche de 1, plus il est vraisemblable que ce point appartienne à la classe correspondante.

Les données s'étalant sur plusieurs années sont ensuite mises ensemble en regroupant les cartes de fonction d'appartenance. Ainsi, pour un champ, on obtient trois cartes d'appartenance faisant la synthèse des données acquises sur plusieurs années. Ce regroupement est effectué en appliquant les règles de la logique floue. Ces règles sont l'équivalent, en logique floue, des régles d'opération en logique binaire: "et", "ou" … Pour notre application, trois règles ont été considérées. Le Tableau 1 trace le parallèle entre les opérations booléennes qui sont familières et les opérations en logique floue qui sont correspondantes. En logique binaire, il est impossible de faire la moyenne entre deux valeurs. Cela est possible en logique floue. On peut classer les opérations du Tableau 1 selon l'effet résultant sur la valeur de la fonction d'appartenance. Ainsi, le "et" est une opération qui conserve la valeur la plus faible, la moyenne est une opération intermédiaire alors que le "ou" ne conserve que la valeur maximum. D'une certaine manière, on peut concevoir l'utilisation du "et" comme une approche conservatrice. Par exemple, dans le cas de cartes de rendement, pour un point donné dans le champ, l'application du "et" ne retiendra que la valeur de la fonction d'appartenance correspondant à l'année où celle-ci est la plus faible diminuant, pour ce point, la vraisemblance qu'il appartienne à l'ensemble correspondant (rendement faible, typique ou élevé). C'est donc une approche conservatrice puisqu'elle "disqualifie" plus facilement un point. Il a été possible d'observer ces tendances concernant l'effet du

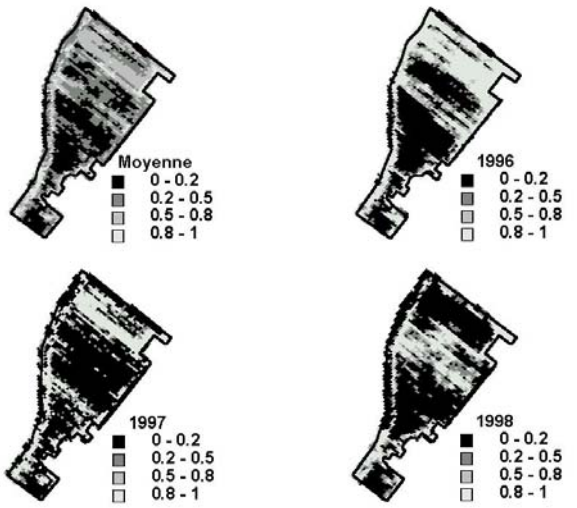


**Fig. 6. Agglomération des fonctions d'appartenance - rendement faible à partir de trois ans de données. Surface de référence = 50%.**

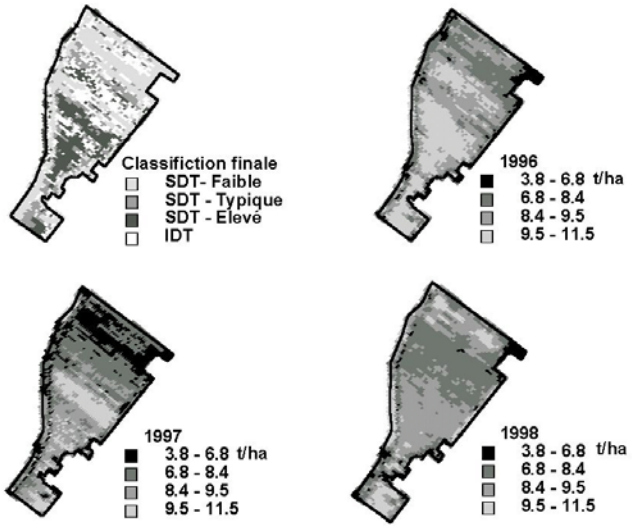


**Fig. 7. Exemple de classification finale. Surface de référence = 50%, seuil = 0.5.**

choix de l'opérateur sur le résultat final. L'opérateur calculant la moyenne a été retenu dans tous les cas. La Fig. 6 montre trois cartes illustrant la fonction d'appartenance à la catégorie des rendements faibles pour trois années et le résultat de la moyenne.

L'étape suivante de la classification consiste en l'application d'un seuil aux trois cartes d'appartenance. Lorsque la valeur de la fonction d'appartenance pour un point donné est inférieur au seuil, ce point est exclu de la classe correspondante. À l'inverse pour des valeurs supérieures ou égales au seuil, le point correspondant est inclus dans la classe correspondante. Le seuil est le deuxième paramètre de la méthode de classification.

La dernière étape consiste à regrouper les données provenant des trois cartes de fonction d'appartenance converties par l'application du seuil en une carte montrant les zones de rendement. La carte finale a quatre catégories (Fig. 7). Un point est assigné à la catégorie "SDT - rendement faible" lorsqu'il a été inclus dans cette classe lors de l'application du seuil et qu'il n'a été inclus dans aucune autre classe. La même logique s'applique pour les catégories "SDT - rendement typique" et "SDT - rendement élevé". Autrement, un point se retrouve dans la catégorie "IDT". La catégorie "IDT" regroupe donc les points qui ont été exclus de toutes les catégories "SDT" lors de l'application du seuil ou qui ont été inclus dans plus d'une catégorie "SDT" (*e.g.* SDT-faible et SDT-élevé). Ces deux situations sont clairement des cas où le rendement a été variable dans le temps.

**Stratégie d'utilisation**

En principe, la méthode peut être utilisée pour combiner des données de rendement obtenues sur des cultures différentes pratiquées sur un même champ sous l'hypothèse que deux cultures qui subissent le même cycle climatique (i.e. saisons de croissance du semis à la récolte qui se superposent) peuvent être mises ensemble lors de la classification. Cette hypothèse doit être évaluée cas par cas en considérant le système de rotation des cultures qui est pratiquée sur la parcelle à l'étude.

Sur le plan pratique, il arrive souvent qu'une parcelle ne puisse pas être considérée comme une entité pour des raisons qui sont liées aux pratiques culturales: semis réparti sur plusieurs jours, utilisation de variétés différentes… Lorsqu'elle est présente, cette variabilité artificielle masque la variabilité naturelle ou intrinsèque que l'on cherche à faire ressortir. La variabilité artificielle a été prise en compte en appliquant la première partie de la méthode de classification (calcul de la fonction d'appartenance pour chaque point) sur des zones de la parcelle ayant subies des traitements jugés uniformes. Les paramètres des fonctions d'appartenance sont alors calculés à partir de l'histogramme construit à partir des données localisées à l'intérieur d'une zone traitée uniformément. Par la suite, la méthode de classification s'applique sans tenir compte de ces zones. La Fig. 8 présente un exemple où trois variétés différentes ont été semées sur une parcelle dont une des variétés affichait un rendement moyen nettement différent des deux autres. En appliquant la classification, l'effet de variété disparaît de la classification finale.

## DISCUSSION

Une analyse statistique (Tukey's Honestly Significant Difference, Dowdy et Wearden 1991) a été réalisée pour vérifier si les données d'une carte de rendement annuel regroupées suivant la classification donne des ensembles de points qui sont significativement différents les uns des autres. Sur une carte annuelle de rendement, tous les points appartenant à une catégorie SDT ont été regroupés et la moyenne du rendement a été calculée pour chaque groupe. Les différences entre ces moyennes ont été soumises au test statistique (Tableau 2). Pour chacune des années, les différences sont hautement significatives ce qui indique que le choix d'utiliser au moins trois classes pour les rendements stables est justifié.

Des diagrammes de dispersion montrant le rendement pour une année donnée à une position donnée dans la parcelle en

**Tableau 2. Différence entre les rendements moyens par classe.**

| Différence (t/ha) | 1996 | 1997 | 1998 |
|---|---|---|---|
| Rendement typique - faible | 1.62* | 1.78* | 0.69* |
| Rendement élevé - typique | 0.68* | 1.46* | 1.01* |

* Différence significative; P<0,01

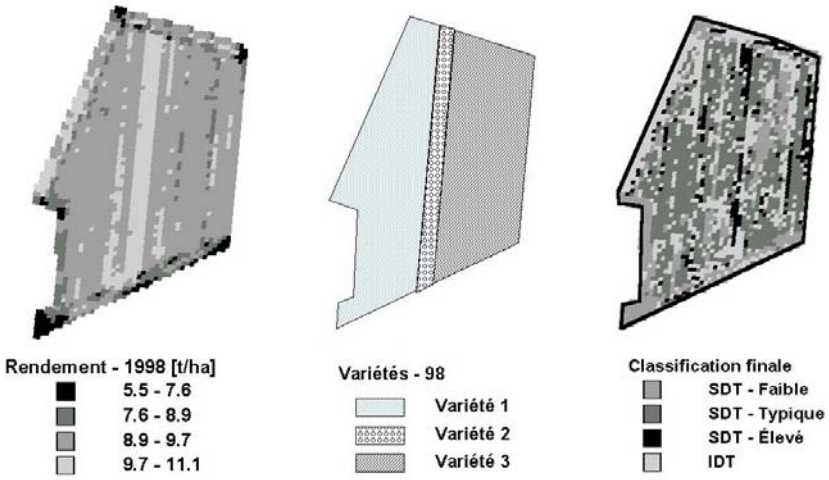


**Fig. 8. Résultat de l'application de la méthode de classification prenant en charge l'effet de variétés. (Maïs - deux ans de données: 1996 et 1998. Des données d'un champ différent ont été utilisées car l'effet de variété n'est pas perceptible dans le champ utilisé pour illustrer la méthode.)**

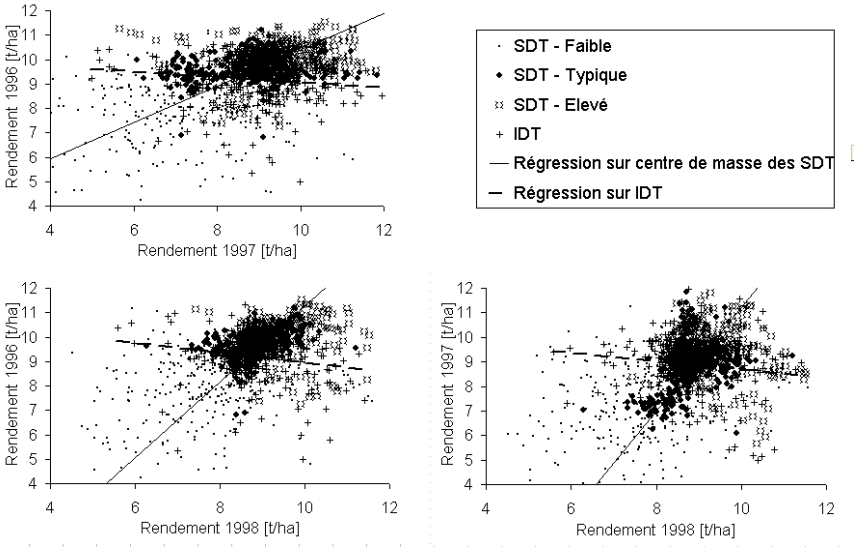


**Fig. 9. Diagramme de dispersion. Surface de référence = 50%, seuil = 0.5.**

fonction du rendement à la même position mais pour une autre année ont été préparés (Fig. 9). Les points ont été regroupés en fonction de la classification de la série temporelle. Les points associés aux classes SDT sont groupés en nuages dont les centres s'alignent le long d'une droite de pente positive ce qui indique que le rendement est corrélé positivement entre les deux années. Les points de la catégorie IDT ont tendances à se loger le long des interfaces entre les classes SDT et les nuages ainsi formés ont une orientation générale le long d'une droite de pente négative. Les nuages de points n'ont pas un contour net et ils se pénètrent les uns les autres. Cela est une indication a posteriori que l'approche par logique floue est indiquée.

La méthode de classification repose sur le choix de deux paramètres : la surface de référence et le seuil. Le choix de ces paramètres a un impact sur la classification finale. Une augmentation de la valeur du seuil pour une surface de référence fixe augmente la proportion des points assignés à la catégorie IDT. Pour une valeur du seuil fixe, une augmentation de la surface de référence augmente le nombre de points classés dans la catégorie SDT-Typique et dans une moindre mesure, le nombre de points dans la catégore IDT. Une augmentation de la surface de référence a aussi pour effet de créer des zones plus homogènes (Panneton et Brouillard 2002).

Il serait désirable de fixer les paramètres en fonction de critères objectifs. La surface de référence est étroitement liée à la définition des catégories SDT. En pratique, ce qui est considéré comme un rendement faible ou élevé dépend de facteurs extérieurs (e.g. prix des grains versus coûts de production). C'est donc à partir de ces facteurs que la surface de référence doit être définie. Par exemple, la frontière entre rendement faible et typique (a sur la Fig. 4) pourrait être basée sur un critère économique comme le rendement en-dessous duquel le retour sur l'investissement est négatif. En définissant a, la position de la barre horizontale sur la Fig. 4 est fixée, ce qui détermine b et la surface de référence.

Le choix du seuil affecte la répartition des points entre les catégories SDT et la catégorie IDT. En retirant du jeu de données, les données correspondant aux points classés dans la catégorie IDT, les coefficients de corrélation linéaire entre les données de rendement annuelles ont augmentés (Fig. 10). Cela montre la pertinence de la méthode pour identifier les points IDT. L'augmentation du coefficient de corrélation pourrait être utilisée pour définir une valeur de seuil approprié. En augmentant graduellement le seuil, les coefficients de corrélation ont tendance à augmenter (Panneton et Brouillard 2002). Le seuil optimum correspondrait alors à la valeur au-delà de laquelle l'augmentation des

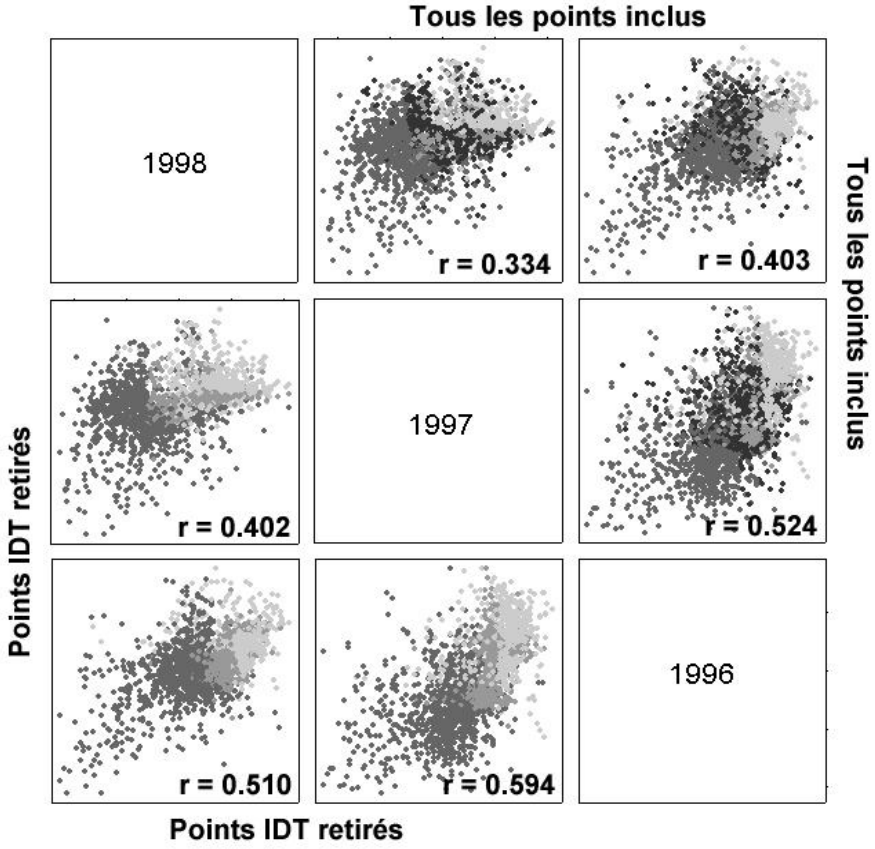


**Fig. 10. Effet des points IDT sur les diagrammes de dispersion et les coefficients de corrélation linéaire, r. Surface de référence = 50%, seuil = 0.5.**

coefficients de corrélation n'est plus statistiquement significative.

La méthode a permis d'isoler les parties de la parcelle où le rendement tendait à être instable dans le temps et de grouper les points où le rendement était plus stable dans le temps en trois catégories : faible, typique et élevé. Les résultats ont montré que les trois catégories SDT sont bien distinctes. La définition précise des catégories dépend de deux paramètres, la surface de référence et le seuil. Des pistes ont été présentées pour raffiner la méthode grâce à une définition rigoureuse de ces paramètres.

La méthode qui a été développée est une méthode de classification supervisée car le résultat de la classification doit se conformer à la définition des classes établie a priori. Cette approche diffère de celle de Lark et Stafford (1998) où la signification des classes doit être découverte a posteriori par des analyses subséquentes. La différence entre ces deux approches réside dans la question qui se pose à l'agronome une fois la classification terminée. Dans le cas de la méthode de Lark et Stafford (1998), la question est: en quoi diffèrent les zones appartenant à des classes différentes. Dans le cas de la méthode qui a été décrite ici, la question est: pourquoi le rendement dans une partie du champ est généralement soit, faible, typique, élevé ou variable. Dans ce cas, l'agronome est tout de suite orienter vers des pistes diagnostiques qui sont propres à chaque catégorie.

## REFERENCES


Blackmore, S., R.J. Godwin, J.C. Taylor, N.D. Cosser, G.A. Wood, R. Earl et S. Knight. 1999. Understanding variability in four fields in the United Kingdom. In *Proceedings of the Fourth International Conference on Precision Agriculture*, 3-18. Madison, WI: American Society of Agronomy.

Blackmore, S. et M. Moore. 1999. Remedial correction of yield data map. *Precision Agriculture* 1: 53-66.

Broos, B., B. Missotten, S. Reybrouck et J. De Baerdemaeker. 1999. Mapping and interpretation of sugar beet yield differences. In *Proceedings of the Fourth International Conference on Precision Agriculture*,171-181. Madison, WI: American Society of Agronomy.

Dowdy, S. et S. Wearden. 1991. *Statistics for Research*, 2nd edition. New York, NY: John Wiley & Sons.

Lark, R.M. et J.V. Stafford. 1996. Classification as a first step in the interpretation of temporal and spatial variability of crop yield. *Aspects of Applied Biology* 46:139-142.

Lark, R.M. et J.V. Stafford. 1998. Information on within-field variability from sequence of yield maps: multivariate classification as a first step of interpretation. *Nutrient Cycling in Agroecosystems* 50:277-281.

Nolin, M.C., A.N. Cambouris et R.R. Simard. 1997. La variabilité des sols: son origine et sa gestion. Dans *Actes du Colloque Nouvelles Technologies en Agriculture,* 35-77. Québec, QC: Ministère de l'Agriculture, des Pêcheries et de l'Alimentation.

Panneton, B. et M. Brouillard. 2002. Use of fuzzy mapping to extract management zones from yield maps. CSAE/SCGR Paper No. 02-207. Mansonville, QC: CSAE/SCGR.

Webster, R. et M.A. Oliver. 1990. *Statistical Methods in Soil Land Resource Survey*. New York, NY: Oxford University Press.

Zhang, Q. et E. Benson. 2000. Fuzzy logic and its application in agriculture - Lecture notes. Department of Agricultural Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.