

# Fusion de données géoréférencées

## Aggregation of georeferenced data

J.N. Paoli<sup>1</sup>

O. Strauss<sup>2</sup>

B. Tisseyre<sup>1</sup>

J-M. Roger<sup>3</sup>

S. Guillaume<sup>3</sup>

<sup>1</sup> Ecole Nationale Supérieure Agronomique - UMR ITAP

<sup>2</sup> Laboratoire d'informatique, de Robotique, et de Microélectronique

<sup>3</sup> Cemagref - UMR ITAP

2 place Viala, 34070 Montpellier cedex 1, paoli, tisseyre@ensam.inra.fr

16 rue Adda, 34090 Montpellier, olivier.strauss@lirmm.fr

361 rue Breton, 34033 Montpellier cedex 1, roger, guillaume@montpellier.cemagref.fr

### Résumé :

Cet article s'intéresse à la description et au traitement d'informations spatialisées. Chaque donnée géoréférencée constitue une source d'information dont la valeur et la position sont définies de manière indépendante. La méthode de traitement proposée utilise l'ensemble des sources pour renseigner une zone de requête quelconque. Elle permet de traiter une large gamme de données, qualitatives ou quantitatives, décrites de manière imprécise et incertaine. Elle peut être appliquée quelle que soit la résolution spatiale des données de départ. La démarche mise en oeuvre s'appuie sur les nombres flous et la théorie des possibilités. La combinaison des informations disponibles met en oeuvre un opérateur d'agrégation basé sur une intégrale de Choquet. Une application de la méthode sur des données de viticulture de précision est également détaillée.

### Mots-clés :

Information spatialisée, théorie des possibilités, intégrale de Choquet, viticulture de précision.

### Abstract:

This article deals with a method to describe and manage spatial knowledge. Each georeferenced data is considered as an information source, whose location and value are independently described. The method proposed in this paper uses the information sources to infer the value of any request zone. This method is able to treat imprecise or uncertain, qualitative and quantitative data. It can be applied whatever the spatial resolution of the available data sets. Our approach is based on fuzzy sets and possibility theory. The aggregation of information sources is provided by an aggregation operator based on the Choquet integral. An application of the method on a precision viticulture data set is also detailed.

### Keywords:

Spatial information, possibility theory, Choquet integral, precision viticulture.

## 1 Introduction

Au cours des dix dernières années, la gestion de données spatialisées a pris une impor-

tance considérable dans de nombreux domaines d'application : étude des sols, agriculture de précision, etc. De nouvelles technologies, telles que la télédétection, ou le développement de capteurs embarqués sur matériels agricoles, produisent un volume de plus en plus important de données géoréférencées (DG).

Une DG se définit comme l'instanciation d'une variable  $v$  sur une position géographique. Un ensemble de DG peut être utilisé pour générer une carte mono-variable, ou pour construire un système d'aide à la décision. Quelle que soit l'application envisagée, les DG ne sont généralement pas correctement positionnées. Une première étape dans leur traitement consiste à inférer  $v$  sur les positions géographiques souhaitées, en fonction des DG disponibles. L'objectif est d'une part de superposer les différentes variables de manière à pouvoir les combiner, et d'autre part de renseigner les zones présentant un intérêt. La définition de ces zones, que nous appelons zones de requête, dépend uniquement des objectifs de l'étude. Elle est indépendante de la variable étudiée et s'affranchit complètement des problèmes de résolution.

Les variables étudiées (dont un exemple est fourni figure 1) présentent une hétérogénéité de résolution et de nature. Une même zone d'étude peut contenir des échantillons dont la taille varie de 1 à 1000 (3 analyses de sol pour 3000

points de rendement). Ces valeurs peuvent être quantitatives ou qualitatives. Leur position peut être déterminée grâce à un système de mesure ou à la connaissance d'un expert. Les données générées sont généralement imprécises et incertaines. En effet :

- les capteurs utilisés pour la mesure comme pour la localisation sont imprécis,
- un expert utilise des termes linguistiques imprécis et incertains pour instancier une valeur et ne possède pas toute l'information nécessaire à sa localisation,
- la continuité spatiale de la plupart des phénomènes étudiés est incompatible avec la définition de frontières précises.

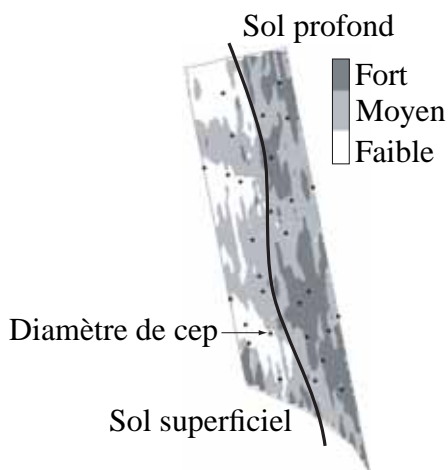


FIG. 1 – Carte de rendement sur vigne et données complémentaires

Cet article présente une nouvelle méthode de fusion dont le but est d'inférer la valeur d'une variable sur une zone de requête quelconque à partir de l'ensemble des données géoréférencées. Elle peut traiter des données qualitatives comme quantitatives et prendre en compte les imprécisions et incertitudes précédemment définies.

Cette méthode interprète les données de manière experte. La mesure est traduite sous forme de termes linguistiques caractérisés par des degrés de confiance. L'inférence au sein d'une zone de requête est la sélection des termes

linguistiques compatibles avec les données présentes. L'absence de données, ou la présence de données trop bruitées se traduit par une indétermination de la zone.

La section 2 décrit notre problématique, et introduit les principaux concepts utilisés par notre méthode. La section 3 présente les aspects théoriques de notre approche. Notre méthode a été appliquée sur des données de viticulture de précision. Les résultats obtenus sont présentés dans la section 4.

## 2 Problématique

L'instanciation d'une variable sur une position géographique passe usuellement par l'interpolation des données géoréférencées. Parmi l'ensemble des méthodes d'interpolation disponibles, les plus couramment utilisées se basent sur une approche géostatistique ([12, 8, 9, 1, 13]). Ces méthodes s'appliquent parfaitement aux données numériques à forte résolution. Mais elles ne permettent ni de prendre en compte des données qualitatives, ni d'introduire des imprécisions et des incertitudes sur les données. C'est la raison pour laquelle ce travail s'intéresse à d'autres méthodes de représentation et de fusion de l'information spatialisée.

Deux approches distinctes ont actuellement été envisagées pour la représentation de données imprécises et incertaines.

La première approche traite les données géoréférencées comme des données usuelles. Les ensembles flous et la théorie des possibilités ([19, 5]) sont utilisés pour prendre en compte le caractère imprécis et incertain de la mesure. Ces méthodes sont employées pour décrire et combiner un ensemble de variablesinstanciées sur les mêmes positions ([11, 3]). Le référentiel de la mesure est partitionné en sous ensembles flous, décrivant des termes linguistiques (par exemple : faible, moyen, fort). La mesure est décrite par un sous ensemble flou. Elle permet de calculer pour chacun des termes

précédemment définis des degrés de confiance, nommés possibilité et nécessité. L'intérêt de cette démarche est double. D'une part, elle prend en compte le caractère imprécis de la mesure. D'autre part, elle permet de se ramener, quelle que soit la variable de départ, à un référentiel qualitatif.

La seconde approche s'intéresse exclusivement aux problèmes de localisation. Elle s'appuie sur la définition d'objets spatiaux délimités par des frontières imprécises (une montagne, une forêt, etc). Ces objets sont principalement décrits par des modèles statistiques ou par des ensembles flous ([2]). Schneider définit différents types de sous ensembles flous géographiques, qu'il désigne par le terme de régions floues ([16]).

Afin de parvenir à une description complète des imprécisions et incertitudes associées aux DG, une combinaison de ces deux approches est proposée à la section 3.1. La position d'une DG est désignée par une région floue, alors que la mesure est décrite par une approche possibiliste.

La zone de requête est ensuite caractérisée par des estimations supérieures de possibilité (et inférieures de nécessité). Elles sont obtenues par l'agrégation des degrés fournis par chaque DG. L'importance d'une DG dans le processus d'agrégation dépend de sa position par rapport à la zone de requête. Cependant, au regard de l'ensemble des valeurs disponibles pour une variable, nous souhaitons identifier et éliminer d'éventuelles données aberrantes. Cela signifie qu'une DG ne constitue pas une information indépendante. Il existe des synergies positives ou négatives entre les données décrivant une même variable.

Parmi tous les opérateurs d'agrégation disponibles, ceux qui prennent en compte ces interactions complexes utilisent des mesures floues ([10]). Leur utilisation a été proposée par Sugeno en 1974 ([17]). Une mesure floue est un type particulier de fonction qui évalue l'importance d'un ensemble de critères dans un processus d'agrégation. Les opérateurs d'agrégation

prenant en compte les mesures floues sont les intégrales floues ([17, 7]). La section 3.2 décrit un opérateur de fusion basé sur l'intégrale de Choquet. Cette intégrale floue a été définie par Murofushi et Sugeno en 1989 ([14]), d'après un concept introduit par Choquet en 1953 ([4]).

Une application est présentée à la section 4. Il s'agit d'abord d'une cartographie des zones de rendement faible, moyen, et fort, d'une parcelle de vigne. Nous utilisons ensuite les cartes obtenues pour positionner des points d'échantillonnage sur lesquels peuvent être réalisées des mesures complémentaires. La mesure de rendement étant quantitative, le résultat obtenu peut être comparé avec celui d'une interpolation spatiale. Notre objectif est double. Il s'agit de montrer d'une part que les résultats de notre approche sont cohérents, et d'autre part que nos cartes permettent de mieux positionner ces points d'échantillonnage, dans la mesure où elles permettent de distinguer les zones où un niveau de rendement est certain des zones où il est simplement possible.

### 3 Théorie

Cette partie s'intéresse d'abord à la description d'une DG, ensuite à la fusion d'un ensemble de DG sur une zone de requête, et enfin aux propriétés de l'opérateur d'agrégation choisi.

#### 3.1 Définitions

**Référentiel.** Une information géoréférencée est définie simultanément par deux attributs : position et mesure. L'attribut position prend sa valeur sur un référentiel de localisation  $L$  (par exemple  $L = \mathbb{R}^{+2}$ ). L'attribut mesure est associé à un ensemble de référence  $M$  (par exemple :  $M = [0, 20]$ ).

Soit  $\mathcal{P}(M)$  une partition de  $M$ , telle que :

$$\mathcal{P}(M) = \{(A_i), i \in [1, \dots, p]\} \quad (1)$$

$\mathcal{P}(M)$  désigne l'ensemble des labels que peut prendre l'attribut mesure ( $A_1 =$  faible,  $A_2 =$  moyen,  $A_3 =$  fort).

**Élément d'information géoréférencé.** Un élément d'information géoréférencé  $I$  est une entité discrète qui restreint  $L \times \mathcal{P}(M)$ . Cette restriction est définie de la manière suivante :

$$I = (P(I), \Pi(I, A_i), N(I, A_i)) \quad (2)$$

$$i \in [1, \dots, p]$$

$P(I)$  est un sous ensemble flou de  $L$ .

$\Pi(I, A_i)$  et  $N(I, A_i)$  sont les possibilité et nécessité de  $A_i$ .

**Objectif de la fusion.** Soit  $R$  une zone de requête quelconque.  $R$  est un élément d'information dont la position est décrite par un sous ensemble flou  $P(R)$ .

Notre objectif est de proposer des estimations de  $\Pi(R, A_i)$  et  $N(R, A_i)$ , à partir d'un ensemble  $S$  d'éléments d'information (appelés sources d'information).

$$S = \{I_k, k \in [1, \dots, n]\} \quad (3)$$

### 3.2 Estimation des possibilités et nécessités pour la zone de requête

**Estimation des possibilités.** L'estimation de  $\Pi(R, A_i)$  est basée sur une intégrale de Choquet :

$$\widehat{\Pi}(R, A_i) = \quad (4)$$

$$\sum_{i=1}^n \Pi(I_{(k)}, A_i) [F(T_{(k)}) - F(T_{(k+1)})]$$

$$\Pi(I_{(1)}, A_i) \leq \dots \leq \Pi(I_{(n)}, A_i)$$

$$T_{(k)} = \{I_{(k)}, \dots, I_{(n)}\}, k \in [1, n] \quad (5)$$

$$T_{(n+1)} = \phi \quad (6)$$

$F$  est une mesure floue définie de la manière suivante :

$$2^S \rightarrow [0, 1] \quad (7)$$

$$T \rightarrow F(T)$$

$$F(\phi) = 0 \quad (8)$$

$$T_2 \subseteq T_1 \Leftrightarrow F(T_2) \leq F(T_1) \quad (9)$$

**Estimation des nécessités.** Soit  $\bar{A}_i$  le complémentaire de  $A_i$ . L'estimation de la possibilité de  $\bar{A}_i$  est obtenue comme précédemment par une intégrale de Choquet.

$$\widehat{\Pi}(R, \bar{A}_i) = \quad (10)$$

$$\sum_{i=1}^n \Pi(I_{(k)}, \bar{A}_i) [F(T_{(k)}) - F(T_{(k+1)})]$$

L'estimation de la nécessité de  $A_i$  est déduite de la relation  $\widehat{N}(R, A_i) = 1 - \widehat{\Pi}(R, \bar{A}_i)$

**Définition de la mesure floue.** Le choix de  $F$  détermine les propriétés de l'estimateur défini par l'intégrale de Choquet. Nous souhaitons en premier lieu réaliser une estimation supérieure de possibilité. Une statistique agrégeant les valeurs les plus hautes peut convenir. Toutefois, des données aberrantes doivent être éliminées. Il nous faut donc configurer un opérateur statistique avec rejet.

Lorsque chacune des sources peut être utilisée pour renseigner la zone de requête, notre problème se ramène à une statistique d'ordre (L-statistique). Nous nous intéressons à un opérateur de type moyenne tronquée. Murfushi et Sugeno ont montré que ce type d'opérateur pouvait être obtenu à partir de l'intégrale de Choquet, en définissant une mesure floue  $F$  basée sur les cardinalités des ensembles  $T_{(k)}$  ([15]).

Nous proposons de généraliser l'approche par moyenne tronquée aux cas où les sources ont une importance variable dans le processus d'agrégation. Pour cela, nous remplaçons la cardinalité par une quantité d'interaction  $Q$  définie de la manière suivante :

$$\alpha_k = \Pi(P(R), P(I_k)) \quad (11)$$

$$Q_{(k)} = \sum_{l=k}^n \alpha_l \quad (12)$$

$\alpha_k$  représente une possibilité d'interaction entre la zone de requête  $R$  et une unique source

d'information  $I_k$ . Les nécessités d'interaction  $N(P(R), P(I_k))$  et  $N(P(I_k), P(R))$  ne sont pas évaluées. Leur utilité et leur prise en compte éventuelle pourront faire l'objet d'une étude ultérieure.

Nous réalisons ensuite une moyenne tronquée de l'ensemble des  $I_{(k)}$  tels que :

$$Q^\beta - \Delta Q \leq Q_{(k)} \leq Q^\beta \quad (13)$$

$$Q^\beta = \beta Q_{(1)}, \beta \in [0, 1] \quad (14)$$

$$\Delta Q \geq 0$$

La fonction  $F$  est représentée figure 2.

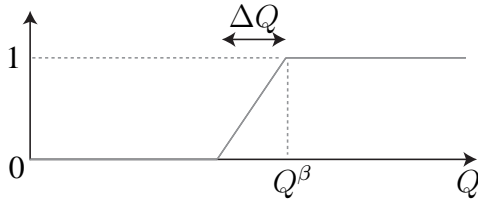


FIG. 2 – Forme générale de la fonction  $f$

$Q_{(1)}$  représente la quantité d'interaction totale entre la zone de requête et l'ensemble des sources.  $\beta$  conditionne le taux de rejet. Les valeurs élevées de possibilité, associées à une faible quantité d'interaction sont considérées comme atypiques. Pour conserver une estimation supérieure de possibilité, il faut rejeter moins de la moitié des informations présentes, et donc imposer  $\beta < 0,5$ .

$\Delta Q$  permet de sélectionner parmi les valeurs restantes de possibilité celles qui seront combinées. L'augmentation de  $\Delta Q$  permet d'incorporer dans l'estimation des valeurs de plus en plus faibles de possibilité, associées à une quantité d'interaction de plus en plus importante. Nous considérons que la quantité d'interaction nécessaire à l'établissement d'une valeur est  $\Delta Q = 1$ . Cela signifie qu'une source peut imposer seule sa valeur si et seulement si  $\alpha_k = 1$ , c'est à dire si elle intersecte totalement la zone de requête.

**Comportement de l'estimateur lorsque la quantité d'interaction est faible.** Dans les sections précédentes, il a été établi que  $F(\phi) = 0$  et  $\Delta Q = 1$ .

Ces assertions nous imposent  $Q^\beta \geq 1$ . Ceci signifie que lorsque la quantité d'information est faible, le taux de rejet doit être limité de manière à garantir une quantité d'interaction minimale entre l'ensemble des sources et la zone de requête.

Si  $Q_{(1)} < 1$  alors  $F(S) < 1$ ,  $S$  étant l'ensemble des sources. L'intégrale de Choquet n'est plus normalisée à 1. En conséquence, les estimations de possibilité ne sont plus normalisées à 1. On peut uniquement garantir l'inégalité :

$$\hat{\Pi}(R, A_i) + \hat{\Pi}(R, \bar{A}_i) \geq f(Q_{(1)}) \quad (15)$$

Une renormalisation peut toutefois être effectuée de la manière suivante :

$$\tilde{\Pi}(R, A_i) = \quad (16)$$

$$\max(\hat{\Pi}(R, A_i), 1 - \hat{\Pi}(R, \bar{A}_i))$$

$$\tilde{\Pi}(R, \bar{A}_i) = \quad (17)$$

$$\max(\hat{\Pi}(R, \bar{A}_i), 1 - \hat{\Pi}(R, A_i))$$

### 3.3 Propriétés

Nous présentons ici un résumé des propriétés les plus importantes de notre estimateur

**Continuité spatiale.** L'utilisation de sous-ensembles flous du référentiel géographique garantit qu'une distorsion continue de  $P(R)$  entraîne une variation continue de  $Q^\beta$ , et par conséquence une variation continue de  $\Pi(R, A_i)$  et  $N(R, A_i)$ .

**Exactitude.** Soit  $R$  une zone de requête complètement décrite par une unique source d'information  $I_1$ . Les estimations de  $\Pi(R, A_i)$  et  $N(R, A_i)$  sont exactes. Les valeurs calculées sont égales aux valeurs observées.

$$\hat{\Pi}(R, A_i) = \Pi(I_1, A_i)$$

$$\hat{N}(R, A_i) = N(I_1, A_i)$$

**Gestion des conflits entre sources .** Une source d'information atypique est éliminée. En cas de désaccord entre plusieurs groupes importants de sources, la zone de requête est instantanée par plusieurs valeurs possibles mais non nécessaires.

**Incomplétude des données .** Une unique source d'information suffit à décrire une zone de requête. Si les interactions entre la zone de requête et les sources d'information sont insuffisantes, les possibilités et nécessités estimées diminuent au point de n'être plus normalisées : aucune valeur n'est possible. Une renormalisation des valeurs entraîne une situation analogue à celle décrite en cas de conflit : toutes les valeurs sont possibles mais non nécessaires.

**Paramétrage .** Le principal paramètre de l'estimateur est  $\beta$ , qui conditionne le taux de rejet. S'il est trop élevé, des sources atypiques sont prises en compte. Ce qui entraîne un conflit entre sources, et donc une indétermination de la zone de requête. Un paramètre trop faible entraîne au contraire un rejet d'informations cohérentes, qui pourraient aider à l'inférence de la valeur de la zone de requête.

## 4 Application

### 4.1 Matériel et méthode

**Données expérimentales .** La méthode précédemment décrite est mise en oeuvre sur des données de Viticulture de Précision. Le vignoble expérimental est une parcelle de Merlot située en Navarre (domaine Julian Chivite, Espagne). Elle couvre une surface de 1,5 hectares. Le rendement de cette parcelle a été mesuré et cartographié en 2002 au moyen d'un capteur embarqué sur une machine à vendanger (société Pellenc S.A) munie d'un dGPS. La résolution des données obtenues avec ce dispositif est élevée (4000 points/ha). De précédents travaux ont permis d'établir et de caractériser les imprécisions attendues pour le système de positionnement et le capteur de

rendement ([18, 6]). Pour l'aspect spatial, nous considérons qu'un point de mesure peut être pris en compte dans un rayon d'un mètre, et ne peut être ignoré dans un rayon de 50 cm, autour de la position donnée par le dGPS (figure 3.a). Concernant la mesure, une valeur comprise dans un intervalle de plus ou moins 2% autour de la mesure peut être acceptée, et une valeur située dans un intervalle de plus ou moins 1% autour de la mesure ne peut être rejetée (figure 3.b).

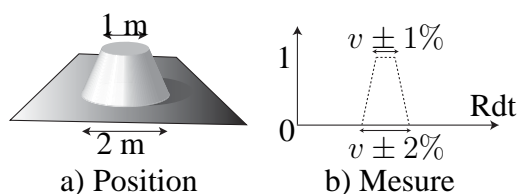


FIG. 3 – description d'une donnée de rendement

**Objectifs de l'étude .** Ce jeu de données est utilisé pour localiser les zones de production dans lesquelles le rendement est conforme, supérieur, ou inférieur aux objectifs de production définis par un cahier des charges (imposé par un décret d'appellation ou par le responsable du domaine). L'objectif est de positionner à partir des cartes obtenues des mesures complémentaires telles que des analyses de sol. Ces mesures permettront de caractériser les différentes zones, de diagnostiquer les problèmes, et de les corriger. Le rendement est décrit par 3 niveaux qualitatifs (faible, correct, fort), définis figure 4.

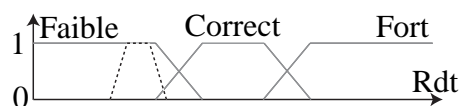


FIG. 4 – Partition du référentiel du rendement

**Définition des zones de requête .** La parcelle est partitionnée en un millier d'unités indépendantes, constituant des zones de requête. Ces zones de requête sont définies de manière imprécise pour tenir compte de

la continuité spatiale du phénomène étudié, et du positionnement imprécis de l'opérateur qui sera chargé de réaliser les prélèvements complémentaires.

Le partitionnement (figure 5) est défini par  $\sup_i \{P(R_i)\} = 1$ ,  $\{R_1, \dots, R_m\}$  étant l'ensemble des zones de requête.

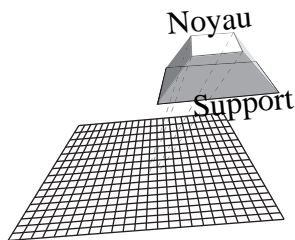


FIG. 5 – Partitionnement de la parcelle

**Opérateur d'agrégation .** Il nous faut déterminer le taux de rejet permettant à notre opérateur d'agrégation d'éliminer les valeurs atypiques de possibilité. Nous décidons de paramétrer l'intégrale de Choquet avec  $\beta = 0,25$ . Nous rejetons donc moins du quart des informations disponibles.

**Protocole de test .** Les cartes obtenues sont jugées selon deux critères. Elles doivent premièrement être cohérentes avec une carte krigée, obtenue selon la méthode de référence en viticulture définie par Bramley en 2001 ([1]). Elles doivent deuxièmement permettre un meilleur positionnement des prélèvements complémentaires.

## 4.2 Résultats et discussion

**Cartes de possibilité (figure 6) .** Ces cartes sont obtenues par l'union des zones de requête précédemment définies. Chaque carte est divisée en 3 zones (gris foncé, gris clair, blanc). La valeur considérée est nécessaire en zone foncée ( $N(R, A_i) = 1$ ), et n'est pas impossible en zone claire ( $\Pi(R, A_i) > 0$ ).

La comparaison entre les cartes de possibilité et la carte krigée (présentée figure 7) est

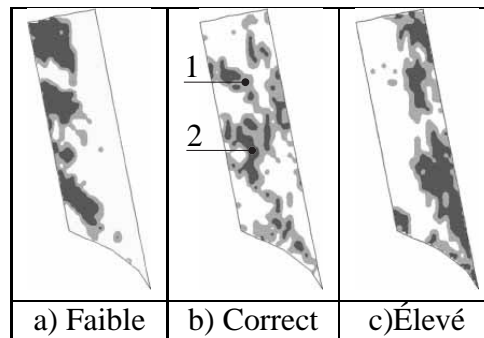


FIG. 6 – Cartes de possibilité

cohérente. les rendements faibles, moyens et forts sont localisés au mêmes endroits.

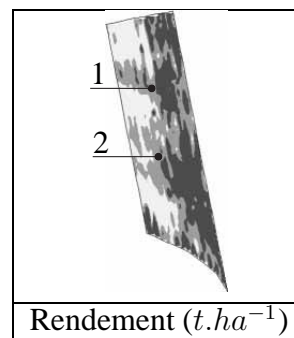


FIG. 7 – Carte obtenue par krigage ordinaire

**Utilisation de l'information .** La méthode présentée permet de définir les zones de la parcelles qui correspondent nécessairement à des rendements faibles, moyens ou fort en prenant en compte l'imprécision du système de mesure et du système de localisation. Cette information constitue une aide à la décision précieuse pour définir un échantillonnage du sol orienté permettant d'effectuer des mesures complémentaires correspondant nécessairement à des rendements faibles, moyen ou fort. Cette application relativement simple reste délicate à effectuer avec une carte krigée (figure 7). En effet, sur ce type de carte à haute résolution, le krigage effectue un lissage des données sur un voisinage qui dépend du modèle choisi, mais ne prend pas en compte l'imprécision du positionnement. L'utilisation d'une telle carte reste difficile



(voire inappropriée) pour positionner un point d'échantillonnage sur une zone dont on souhaite être certain de la valeur. Dans notre exemple, c'est le cas de la zone de rendement moyen qui correspond pour partie à une moyenne de rendements faibles et forts. A titre d'exemple, deux points ont été positionnées sur la carte krigée (figure 7), puis reportés sur la carte de possibilité (figure 6.b). Les deux points sont tous deux localisés en zone moyenne sur la carte krigée. Mais la carte de possibilité montre que le point n°1 n'est pas placé dans une zone nécessairement correcte.

## 5 Conclusion

Dans cet article, nous avons présenté une méthode permettant de traiter une large gamme de données géoréférencées : imprécises et incertaines, qualitatives ou quantitatives, quelle que soit leur résolution spatiale. Un opérateur de fusion paramétrable a été défini pour la combinaison des informations spatiales. L'application a montré la pertinence de cet opérateur sur le jeu de données utilisé. A la suite de cette première étude, un certain nombre de questions restent ouvertes. Tout d'abord, l'interaction entre un ensemble de sources d'information et une zone de requête pourrait être définie de manière plus complète. Ensuite, les problèmes de normalisation, en présence de peu d'information, pourront être mieux cernés. Enfin, une comparaison plus poussée avec les méthodes usuelles d'interpolation spatiale pourra être effectuée.

## Références

- [1] R.G.V Bramley. A protocol for winegrape yield maps. In S. Blackmore and G. Grenier, editors, *Proceedings of Third European Conference on Precision Agriculture*, volume 2, pages 767–173. Agro Montpellier, 2001.
- [2] P.A. Burrough. Natural objects with indeterminate boundaries. In *Geographic objects with Indeterminate Boundaries*, volume 2 of *GISDATA*, pages 3–28. Taylor & Francis, 1996.
- [3] D.R. Cazemier, P. Lagacherie, and R. Martin-Clouaire. A possibility theory approach for estimating available water capacity from imprecise information contained in soil databases. *Geoderma*, 103(1-2) :113–132, 2001.
- [4] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5 :131–295, 1953.
- [5] D. Dubois and H. Prade. *Possibility theory*. Plenum Press, New-York, 1988.
- [6] M. Ehrl, W. Stempfhuber, H. Auernhammer, and M. Demmel. Quality assessment of agricultural positioning and communication system. In J. Stafford and A. Werner, editors, *Proceedings of Fourth European Conference on Precision Agriculture*, volume 1, pages 205–210, Berlin, 2003. Wageningen Academic.
- [7] M. Grabisch, T. Murofushi, and M. Sugeno. *Fuzzy measures and integrals - Theory and applications*. Studies in Fuzziness and soft computing. Physica-Verlag Heidelberg, New York, 2000.
- [8] A.G. Journel and C.J. Huijbregts. *Mining Geostatistics*. Academic Press, London, 1978.
- [9] R. Kerry and M.A. Oliver. Comparing spatial structures in soil properties and ancillary data by using variograms. In S. Blackmore and G. Grenier, editors, *Proceedings of Third European Conference on Precision Agriculture*, volume 1, pages 413–418. Agro Montpellier, 2001.
- [10] J-L. Marichal. *Aggregation operators for multicriteria decision aid*. PhD thesis, Université de Liège, Faculté des sciences, 1999.
- [11] R. Martin-Clouaire, D.R. Cazemier, and P. Lagacherie. Representing and processing uncertain soil information for mapping soil hydrological properties. *Computers and Electronics in Agriculture*, 29(1-2) :41–57, 2000.
- [12] G Matheron. *La théorie des variables régionalisées et ses applications.*, volume 5 of *Les cahiers du Centre de Morphologie Mathématique de Fontainebleau*. Ecole Nationale Supérieure des Mines de Paris, 1970.
- [13] A.B. McBratney, M.L. Mendonça Santos, and B. Minasny. On digital soil mapping. *Geoderma*, 117(1-2) :3–52, 2003.
- [14] T. Murofushi and M. Sugeno. An interpretation of fuzzy measure and the choquet integral. *Fuzzy sets and systems*, 29 :201–227, 1989.
- [15] T. Murofushi and M. Sugeno. Some quantities represented by the choquet integral. *Fuzzy sets and systems*, 56 :229–235, 1993.
- [16] M. Schneider. Uncertainty management for spatial data in databases : Fuzzy spatial data types. In *6th Int. Symp. on Advances in Spatial Databases (SSD)*, pages 330–351. Springer Verlag, 1999.
- [17] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokio institute of technology, 1974.
- [18] B. Tisseyre, C. Mazzoni, N. Ardoin, and C. Clipet. Yield and harvest quality measurement in precision viticulture - application for a selective vintage. In S. Blackmore and G. Grenier, editors, *Proceedings of Third European Conference on Precision Agriculture*, volume 1, pages 133–138. Agro Montpellier, 2001.
- [19] R. R. Yager, S. Ovchinnikov, R. M. Tong, and H. T. Nguyen. *Fuzzy Sets and Applications : Selected Papers by L.A. Zadeh*. John Wiley & Sons, New York, 1987.