

# ESPACE, ENTROPIE, ENVIRONNEMENT

**Michel Phipps.** *Département de Géographie, Univ. d'Ottawa, CP 450, Succ. A, Ottawa, ON, Canada, K1N 6N5*

*Courriel : mhipps@uottawa.ca*

---

*RÉSUMÉ. L'avènement de la modélisation à base individuelle et le développement des systèmes d'information géographiques soulèvent un nouvel intérêt pour les applications géographiques de la théorie de l'information. Un algorithme conçu pour l'analyse des distributions spatiales à variables discrètes est présenté. Il consiste à réduire progressivement l'entropie du système en utilisant l'information fournie par une combinaison de variables décrivant ce système. Les résultats de cette démarche font apparaître une redondance qui mesure le niveau d'ordre du système. Une application à la distribution des types de couverts forestiers au Manitoba (Canada) est présentée.*

*ABSTRACT. The advent of individual-based modeling and the development of geographical information systems are now fostering a new interest for applications of information theory in geography. An algorithm designed for the analysis of discrete variables distributions is discussed in this paper. It seeks a progressive reduction of the system's entropy using information provided by a particular combination of variables that describe this system. The results of this procedure yield a redundancy which measures the system level of order. An application to the spatial distribution of forest cover types in Manitoba (Canada) is shown.*

*MOTS CLÉS : Distributions spatiales discrètes, Entropie, Négentropie, Redondance, Organisation, Algorithme prédictif, Couverts forestiers Manitoba.*

*KEY WORDS : Discrete variable distributions, Entropy, Negentropy, Redundancy, Organization, Predictive algorithm, Forest covers Manitoba.*

---

## 1 Introduction

Au cours des décennies soixante et soixante-dix, on a assisté à l'éclosion d'une abondante littérature sur les applications géographiques de la théorie de l'information. Des revues telles que *Geographical Analysis* donnèrent à l'époque un large écho à cette activité de recherche portant généralement sur la distribution de phénomènes quantitatifs dans une grille d'unités spatiales (population, flux de transport). L'approche quasi exclusive privilégiée alors consistait à simuler une distribution en fonction de contraintes spécifiques (minimisation du coût global de transport, par exemple). L'outil de cette optimisation reposait sur le concept de *maximisation de l'entropie* (voir par exemple Wilson, 1970 et Webber, 1977). En dépit de succès notables, l'engouement suscité par ce genre d'applications s'atténua rapidement et on trouverait certainement très peu de titres publiés en ce domaine au cours des années 90<sup>1</sup>

Pourtant, le temps semble venu aujourd'hui de réexaminer l'intérêt de ce cadre théorique. On verra à ceci plusieurs raisons. L'explosion des capacités de calcul, en vitesse et en volume, a ouvert la voie à de nouvelles formes de modélisation qui eurent naguère passé pour utopiques. C'est le cas de la modélisation à base individuelle (DeAngelis et Rose, 1992) qui permet de bâtir des modèles reposant sur l'individualité de chaque élément du système, évitant ainsi d'écraser ces individus dans des agrégats commodes parce que dénombrables mais parfois dépourvus de sens. Dans le même esprit, cette modélisation à base individuelle autorise une véritable prise en compte des qualités attachées à ces individus. Notons en outre que de tels modèles peuvent mettre à profit l'abondante information contenue dans les systèmes d'information à référence spatiale. Enfin, et c'est là certainement l'argument le plus fort plaidant en faveur de ce regain d'intérêt, la théorie de l'information offre des outils d'une grande puissance opératoire dans les problèmes d'analyse spatiale, puissance qui ne semble guère avoir été soulignée jusqu'à maintenant.

Dans cet article nous définirons d'abord un système spatial distribué à variables discrètes, et rappellerons quelques concepts de la théorie de l'information. Puis nous présenterons un algorithme prédictif applicable à ce type de système. Une application à la distribution des types de couverts forestiers du Manitoba (Canada) sera également présentée.

---

<sup>1</sup> Une recherche portant sur les titres publiés dans 15 revues géographiques entre 1993 et 2000 et affichant des mots clés liés à la théorie de l'information a livré ... une seule référence : voir Antrop, 1998.

## 2 Système spatial distribué à variables discrètes

Considérons un système spatial  $U$  composé d'unités élémentaires (cellules, biotopes, lots, pixels, etc.)  $u_i \in U$  ( $i=1, \dots, N$ ). Chaque unité  $u_i$  est décrite par un ensemble  $V$  de *descripteurs* ou *variables discrètes*  $v_k \in V$  ( $k=1, \dots, l$ ) selon les *états* ou *valeurs*  $a_{ikj} \in A_k$  ( $j=1, \dots, m_k$ ) pris localement par  $v_k$  dans l'unité  $u_i$ . On dira donc, de manière concise, que le système est une application de  $V$  dans l'espace  $U$ .

L'usage des termes de *descripteur*, *variable*, *état* et *valeur* dénote ici que ces descripteurs peuvent être indifféremment des propriétés discrètes de nature qualitative, hiérarchique ou même quantitative attachées aux unités spatiales. Cette précision étant donnée, nous utiliserons systématiquement dans la suite de ce texte les termes de *variable* ou *descripteur* et d'*état*.

On procédera à l'analyse empirique du système  $U$  ainsi défini en dénombrant d'abord les  $n_{kj}$  ( $k=1, \dots, l$  et  $j=1, \dots, m_k$ ) unités où la variable  $v_k$  est dans l'état  $a_{kj}$ ; bien entendu, nous aurons :

$$\sum_{j=1}^{m_k} n_{kj} = N, \quad \forall v_k \in V \quad (1)$$

Ce dénombrement permet ensuite de calculer le nombre de configurations potentielles associées à chacune des variables  $v_k$  de l'ensemble  $V$ . Sous réserve que les unités puissent être considérés comme intrinsèquement non distinguables les uns des autres<sup>2</sup>, ce nombre s'obtient en appliquant la statistique de Fermi-Dirac :

$$W_k = \frac{N!}{\prod_{j=1}^{m_k} n_{kj}!} \quad (2)$$

On voit immédiatement l'intérêt de  $W_k$  en analyse spatiale puisqu'il représente le nombre de configurations spatiales, c'est-à-dire de "*cartes*", pouvant être associées à la variable  $v_k$  ou encore une mesure du potentiel, réel ou virtuel, de *re-configuration* de  $U$  pour cette variable particulière.

Utilisant l'équation de Boltzmann nous obtiendrons finalement une mesure d'entropie pour la variable  $v_k$  :

$$H_k = \frac{1}{N} \log W_k \quad (3)$$

Brillouin (1951) a démontré dans le cas où  $N$  et  $n_{kj}$  sont grands, l'équivalence entre l'équation 3 et l'équation de l'entropie d'information formulée par Shannon :

$$H_k = - \sum_{j=1}^{m_k} p_{kj} \log p_{kj} \quad (4)$$

Nous préférons ici le formalisme Boltzmannien à celui de Shannon en raison de la forte connotation spatiale liée au nombre  $W_k$ . Rappelons à ce propos les autres significations bien connues de l'entropie : incertitude relative à la prédiction de l'état réalisé dans une unité quelconque du système  $U$ , complexité, diversité de ce système et son

<sup>2</sup> Dans l'énoncé de cette clause, le terme *intrinsèquement* signifie que les unités sont non distinguables en ce qui concerne leur nature profonde ce qui n'exclut pas qu'elles puissent éventuellement exprimer des états différents, mais implique que toutes aient la même propension à exprimer chacun de ces états.

contenu d'information <sup>3</sup> mais, aussi, quantité de désordre existant dans ce même système. Nous reviendrons à la question du désordre puisque celui-ci représente *a contrario* un potentiel d'ordre et que le couple antinomique *désordre—ordre* permet de cerner la question fondamentale de la mesure de l'organisation de  $U$ .

Pour l'instant, en utilisant l'équation 3, on pourra évaluer une entropie généralisée du système à partir du nombre total de configurations spatiales que pourrait prendre le système compte tenu de toutes les variables décrivant  $U$  :

$$W_U = \prod_{k=1}^l W_k \text{ et } H_U = \frac{1}{l \times N} \sum_{k=1}^l \log W_k \quad (5)$$

A l'entropie généralisée calculée ci-dessus, on peut également associer une entropie maximum  $H_{max_U}$  sachant que pour chaque variable  $v_k \in V$  le maximum d'entropie est :

$$H_{max_k} = \log m_k \quad (6)$$

valeur réalisée sous la condition d'équi-distribution des unités  $u_i$  dans les états  $a_{kj}$ , c'est-à-dire si :

$$n_{kj} = \frac{N}{m_k}, \forall a_{kj} \in A_k \quad (7)$$

On obtiendra donc cette entropie maximum généralisée suivant :

$$H_{max_U} = \frac{1}{l} \sum_{k=1}^l \log m_k \quad (8)$$

Ces derniers résultats (équations 6-8) sont donnés ici pour mémoire. Leur importance dans les systèmes physiques est considérable puisque la condition donnée par l'équation 7 généralisée à toutes les variables décrivant  $U$  définit l'équilibre thermodynamique à proximité duquel un système évolue vers son entropie maximum. Ici, on doit se poser la question du sens de cette condition dans un système spatial : que signifierait le cas particulier de l'équi-distribution des unités  $u_i$  dans les états  $a_{kj}$  pour l'une quelconque des variables décrivant  $U$  ? Et quel serait *a fortiori* le sens de cette condition généralisée à l'ensemble de ces variables ?

Le développement des systèmes spatiaux résulte de processus de localisation ou de production dont rien ne permet de supposer *a priori* qu'ils puissent conduire à une répartition équitable des unités spatiale entre les états d'une variable. On est donc conduit à penser que la valeur de  $H_{max_U}$  définie par l'équation 8 ne peut guère constituer une référence de quelque signification que ce soit. Tout au plus pourrait-elle représenter un niveau d'incertitude maximum dans lequel se trouverait un observateur dépourvu de toute information sur l'état du système. Et il est clair alors que cette valeur perdrait son intérêt aussitôt qu'une exploration empirique apporterait une information quelconque sur l'état du système.

---

<sup>3</sup> La notion de contenu d'information apparaît clairement si on considère que chacune des  $W_k$  configurations possibles que peut prendre  $U$  forme un arrangement particulier de signes, distinct des autres arrangements ou, en d'autres termes, que cette configuration constitue un message particulier.

Par contre, revenant aux valeurs  $W_U$  et  $H_U$  calculées dans l'équation 5 qui représentent les moyennes empiriques, par variable et par unité élémentaire, du nombre de configurations de  $U$  et son entropie, on doit se poser une nouvelle question beaucoup plus féconde que la précédente. Les deux expressions sont correctement calculées (équation 5) si on tient compte de toutes les variables  $v_k$  décrivant  $U$  mais sans prendre en considération les redondances qui peuvent exister entre ces variables. Or le fait de négliger ces redondances équivaut à surévaluer  $W_U$  et  $H_U$ .

Il existe donc deux valeurs  $(W_U \# W_U)$  et  $(H_U \# H_U)$  qui, compte tenu des redondances, expriment exactement le nombre de configurations possibles de  $U$  et, par conséquent, mesurent son entropie réelle. L'importance de  $(H_U)$  doit être soulignée puisque la différence :

$$R_U = H_U - *H_U \quad (9)$$

donne une mesure des redondances existant au sein de l'ensemble  $V$ , c'est-à-dire des contraintes qu'exercent réciproquement les unes sur les autres les variables  $v_k$ . Bien entendu,  $R_U$  constitue également une mesure du niveau d'organisation de  $U$  compte tenu de la description que nous en avons.

Le calcul de ces nouvelles valeurs  $(W_U)$  et  $(H_U)$  pour l'ensemble des  $l$  variables dépasserait le cadre de cet article. Dans le cas de deux variables  $v_k$  et  $v_{k'}$ , elles sont obtenues en suivant les étapes suivantes :

$$W_{k''\setminus j'} = \frac{n_{j'}!}{\prod_{j''=1}^{m''_{k''}} n_{j'j''}} \quad \text{et} \quad H_{k''\setminus k'} = \prod_{j'=1}^{m'_{k'}} W_{k''\setminus j'} \quad (10)$$

où :

- $W_{k''\setminus j'}$  représente le nombre de configurations possible pour la variable  $v_{k''}$  dans le sous-ensemble des unités élémentaires possédant l'état  $a_{k'j'}$  de la variable  $v_{k'}$  ;
- $n_{j'}$  est le nombre d'unités élémentaires possédant l'état  $a_{k'j'}$  ;
- $n_{j'j''}$  est le nombre d'unités possédant à la fois l'état  $a_{k'j'}$  de  $v_{k'}$  et l'état  $a_{k''j''}$  de  $v_{k''}$ .
- $W_{k''\setminus k'}$  est le nombre de configurations possibles pour la variable  $v_{k''}$  sachant la variable  $v_{k'}$ .

A partir de cette dernière valeur on obtiendra l'entropie correspondante comme on l'a fait précédemment dans l'équation 3 :

$$H_{k''\setminus k'} = \frac{1}{N} \log W_{k''\setminus k'} \quad (11)$$

expression désignée comme l' *entropie conditionnelle* de  $U$  pour la variable  $v_{k''}$  sachant la variable  $v_{k'}$ . Bien entendu, on pourra calculer à partir de cette dernière équation la néguentropie entre  $v_{k'}$  et  $v_{k''}$  comme la différence :

$$I_{k'',k'} = H_{k''} - H_{k''\setminus k'} \quad (12)$$

ainsi que la redondance qui lui est associée :

$$R_{k''\setminus k'} = \frac{I_{k'',k'}}{H_{k''}} \quad (13)$$

On notera en outre que le niveau de signification statistique de la néguentropie peut être testé en utilisant le test du log maximum de vraisemblance (log-likelihood ratio, en anglais) :

$$\xi = 2 \times N \times I_{k'',k'} \approx \chi^2 \text{ avec } (m_{k''} - 1) \times (m_{k'} - 1) \text{ d.l.} \quad (14)$$

Les résultats précédents (équations 10, 11 et 12) ont été acquis en prenant appui sur la variable  $v_{k''}$ , l'autre variable  $v_{k'}$  étant alors prise comme condition. Les calculs pourraient être faits de la même manière en inversant les rôles de ces deux variables. La suite des opérations aurait alors consisté à calculer successivement les valeurs  $W_{k',k''}$  et  $H_{k',k''}$  puis, finalement, la redondance  $R_{k',k''}$ . Mais on notera à cet égard que si la néguentropie est la même dans les deux cas, les valeurs  $W$ ,  $H$  et  $R$  sont différentes selon la voie suivie. On aura donc soin de respecter l'ordre des indices  $k'$  et  $k''$  ainsi que leur position par rapport à la barre oblique qui précise leur rôle respectif.

### 3 Un modèle prédictif multivarié

A partir de ce point, on peut poursuivre l'analyse dans deux directions différentes. Dans une première direction, on peut chercher à évaluer la néguentropie généralisée du système et la redondance qui lui est associée (équation. 9), c'est-à-dire évaluer le niveau d'ordre global du système  $U$  décrit par l'ensemble  $V$ . Cette notion de néguentropie généralisée offre l'intérêt indéniable de saisir une propriété fondamentale du système. Malgré tout, cet intérêt reste tout théorique car cette propriété constitue un point d'aboutissement, une sorte de cul-de-sac méthodologique, dont on perçoit mal les corollaires utiles qu'elle pourrait produire.

La deuxième direction offre au contraire des perspectives opératoires beaucoup plus riches. Elle consiste toujours à évaluer une néguentropie générale du système  $U$  mais, cette fois, par rapport à l'une des variables  $v_{k^*}$  de l'ensemble  $V$ , dite variable-cible. L'analyse se développe alors sous la forme d'un algorithme prédictif visant à prédire les états locaux de cette variable-cible à partir d'une combinaison des états des autres variables de l'ensemble  $V$ . Ces dernières jouent donc le rôle de *prédicteurs* à l'égard de la variable-cible. Transposé dans la terminologie propre aux systèmes d'information géographiques, on peut comprendre cet algorithme comme une tentative de prédire l'une des couches d'information du SIG à l'aide d'une combinaison d'autres couches d'information.

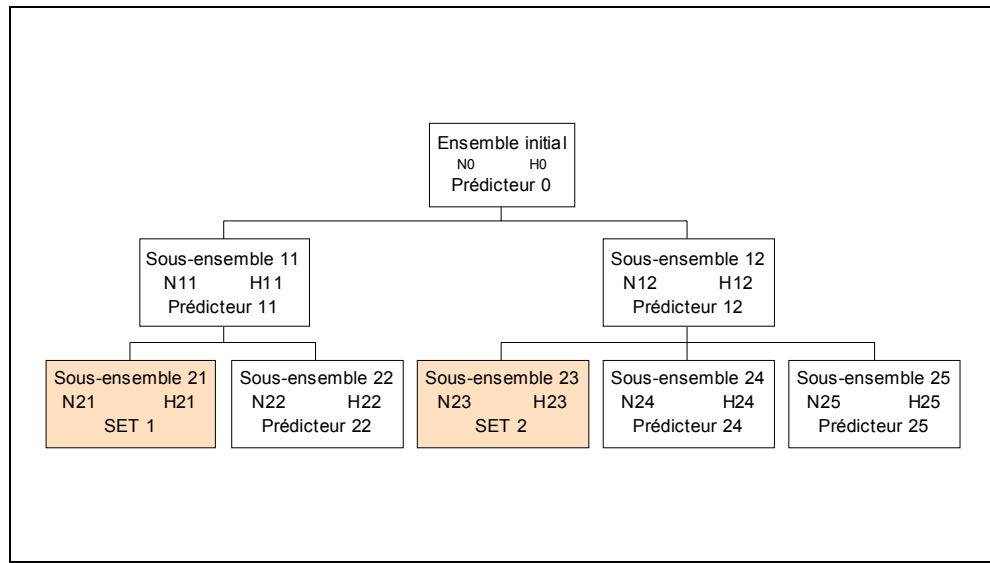
Bien entendu, la combinaison d'états recherchée n'est pas connue. L'objectif de l'algorithme est de la découvrir. Ceci se fait par une procédure de division hiérarchique pas-à-pas maximisant à chaque pas le gain d'information, c'est-à-dire maximisant la néguentropie entre un prédicteur et la variable-cible. La sélection des prédicteurs et l'ordre de leur intervention sont déterminés en fonction de cette néguentropie calculée selon l'équation 12, où la variable-cible  $v_{k^*}$  tient la place de la variable  $v_{k''}$ .

À chaque pas du processus de division, la démarche consistera donc à calculer systématiquement les valeurs de néguentropie :

$$I_{k^*,k} = H_{k^*} - H_{k^* \setminus k} (\forall k \neq k^*) \quad (15)$$

puis à sélectionner le prédicteur en retenant la variable ayant la plus forte néguentropie statistiquement significative. À partir de ce point, les unités  $u_i$  de l'ensemble  $U$  seront distribuées dans  $m_k$  sous-ensembles correspondant aux états du prédicteur  $v_k$  selon les valeurs  $a_{ijk}$  prises localement par cette variable dans les unités du système  $U$ . La suite de la procédure découle naturellement de ce premier pas, chaque sous-ensemble formé étant à son tour traité selon la même démarche.

Figure 1



La figure 1 montre l'amorce du processus de division qui se poursuit jusqu'à ce que toutes les branches de l'arbre soient affectées par l'une des règles d'arrêt. Celles-ci sont au nombre de trois :

- l'entropie du sous-ensemble est inférieure à un seuil au-dessous duquel on considère, à toute fin pratique, que ce sous-ensemble est déterminé;
- le nombre d'unités dans le sous-ensemble est inférieur à un nombre minimum au-dessous duquel la validité du test de signification n'est pas assurée;
- aucune des deux règles précédentes ne s'applique, mais aucune des variables de l'ensemble  $V$  non encore utilisées comme prédicteur n'apporte d'information statistiquement significative.

En résumé, la procédure décrite apparaît donc comme une partition<sup>4</sup> progressive de l'espace  $U$ , selon une combinaison  $C$  d'états des prédicteurs, permettant une ségrégation de plus en plus poussée des états de la variable-cible. Au fur et à mesure de la construction de  $C$ , l'entropie de la variable-cible sachant  $C$  décroît, diminuant ainsi l'incertitude de la prédiction de l'état réalisé dans une unité quelconque  $u_i$ . En même temps, la complexité de  $C$  s'accroît à mesure que de nouveaux sous-ensembles apparaissent. L'entropie attachée à cette combinaison  $C$  mesure cette complexité :

$$H_C = \frac{1}{N} \log \frac{N!}{\prod_{s=1}^v n_s!} \quad (16)$$

où  $v$  est le nombre de sous-ensembles et  $n_s$  le nombre d'unités contenues dans le sous-ensemble  $s$ . La solution combinatoire adoptée pour la recherche de  $C$  confère à celle-ci des propriétés intéressantes sur lesquelles nous reviendrons au cours de la discussion.

#### 4 La distribution des types de couverts forestiers (Manitoba)

À l'exception du triangle NE s'ouvrant sur la baie d'Hudson, le territoire de la province du Manitoba couvre près de 540 000 km<sup>2</sup>. À partir des données de l'inventaire forestier canadien (Grey, 1989), une carte des couvertures

<sup>4</sup> Cette notion de *partition* est à l'origine de la désignation de l'algorithme : PEGASE pour **P**artition d'un ensemble **g**éographique pour l'analyse spatiale **é**cologique

végétales et types forestiers a été réalisée (Jiang, 1996). Trois types de couverts asylvatiques (prairie, terres agricoles, tundra) et sept types de couverts forestiers classés selon leurs espèces dominantes ont été cartographiés sur un réseau spatial cellulaire de près de 6000 cellules.

La question posée dans cette recherche s'articule en deux étapes :

- peut-on prédire la distribution spatiale actuelle des types de couvert et types forestiers à partir de divers prédicteurs environnementaux (climatiques, édaphiques) ?
- une telle prédiction serait-elle suffisamment précise pour prédire une distribution spatiale probable sous l'hypothèse d'un changement climatique donné ?

Dans le cadre de cet article, nous n'examinerons que la première de ces étapes. Une base de donnée comportant pour chacune des cellules la variable-cible couverture et 13 variables environnementales <sup>5</sup> a été soumise à l'algorithme décrit plus haut.

La planche I montre les cartes de cinq des variables de la base de données : les *types de couverts* (variable-cible) (1), les *précipitations annuelles* (2), la *texture des dépôts de surface* (3), la *température moyenne annuelle* (4) et la *température moyenne d'octobre* (5). La planche II montre les résultats cartographiques de l'application de l'algorithme prédictif. Le tableau 1 résume les résultats de l'application. Le processus de division entraîne une chute d'entropie de 1.984 à 0.696 soit une néguentropie de 1.288. La redondance (64.9%) doit être interprétée comme la part d'entropie initiale levée par la combinaison *C* produite par l'algorithme.

**Tableau 1**

Grandeurs	Valeurs	Valeurs ajustées
Entropie initiale	1.98	
Entropie finale	0.70	0.67
Néguentropie	1.29	1.32
Redondance (%)	64.90	66.40
Complexité de C	4.17	4.83

Au cours du processus de division, le nombre de sous-ensembles formés s'accroît déterminant un gain d'information et, par conséquent, une chute d'entropie de la variable-cible ( **O** ) en même temps qu'une complexité croissante de la combinaison *C* ( **□** ) (voir figure 2). Les deux courbes associées à ces symboles sont des fonctions théoriques ajustées sur les valeurs empiriques <sup>6</sup>. Elles permettent de définir les valeurs ajustées figurant dans le tableau 1. Ainsi, l'entropie finale et la néguentropie ajustées représentent-elles respectivement l'asymptote de l'exponentielle négative caractérisant la chute d'entropie et la mesure de l'organisation potentielle du système *U* relativement à l'ensemble *V* de variables décrivant le système. Le degré de précision de la prédiction est attesté à la fois par la valeur de la redondance et la similitude visuelle entre la configuration réelle et les configurations simulées (planche 2).

<sup>5</sup> Les 13 variables sont : 4 variables édaphiques (géologie, dépôts de surface, texture superficielle, régions physiographiques), 6 températures (annuelle, saisonnières, somme des degrés-jour), 3 précipitations (annuelle, hiver, saison de végétation).

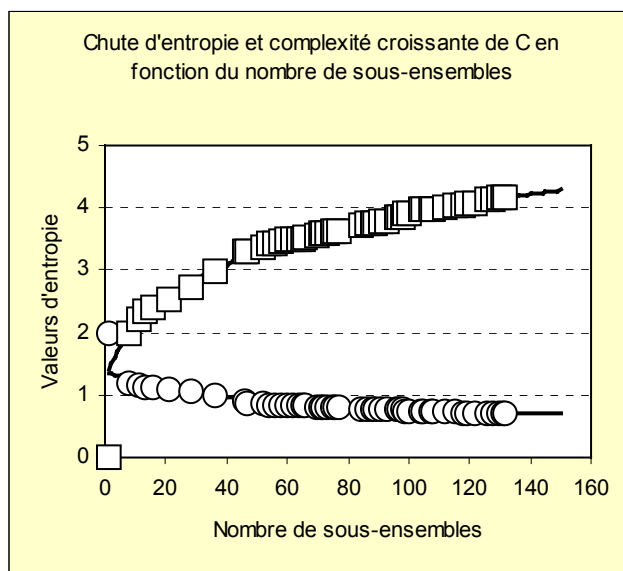
<sup>6</sup> Ces fonctions ajustées sont respectivement la fonction logarithmique :

$$H_C = \log(0.458nse + 3.664) \text{ pour la complexité de } C, \text{ et la fonction exponentielle négative :}$$

$$H_{k^* \setminus C} = e^{-0.022nse - 0.345} + 0.666 \text{ pour la chute d'entropie, fonctions où } nse \text{ est le nombre de sous-ensembles formés au cours du processus de division.}$$



Figure 2



## 5 Discussion

### 5.1 Problématique de la prédiction

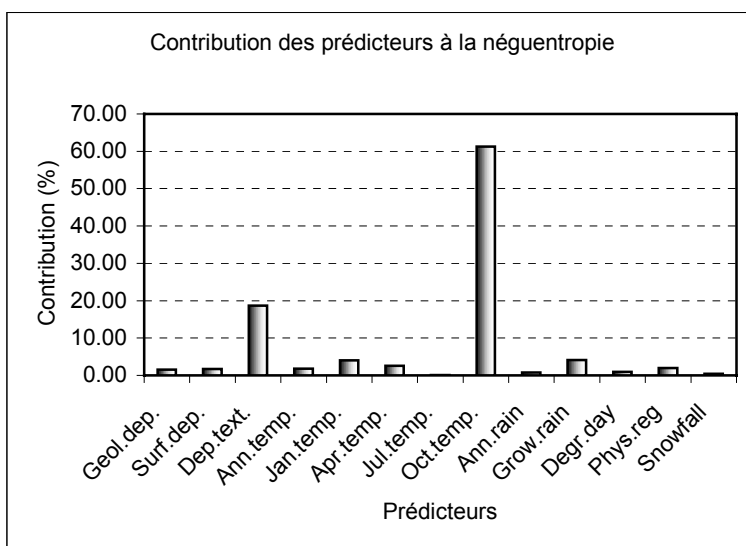
La nature prédictive de l'algorithme décrit ne doit pas induire l'idée que la prédiction de la configuration de la variable-cible serait l'objectif unique et nécessaire de toute application de cette procédure. Bien au contraire. La plupart du temps, l'utilisateur cherche à comprendre les liens entre cette configuration et un phénomène particulier et, par conséquent, à connaître le degré de dépendance de cette configuration à l'égard d'un ensemble de prédicteurs donnés (par exemple, entre une distribution de communautés végétales et des variables environnementales [Phipps, 1981; Pineda *et al.*, 1988; Davis et Dozier, 1989]). Ou, on pourra chercher à comparer deux configurations d'utilisations agricoles du sol quant à leur degré de dépendance à l'égard des caractéristiques pédologiques et édaphiques du territoire (Dumanski *et al.*, 1987; Baudry, 1989). Ou encore tenter de saisir, à travers des analyses diachroniques, des dynamiques d'organisation ou de désorganisation d'un patron par rapport à un ensemble de descripteurs spatiaux (Phipps *et al.*, 1986). De telles problématiques guideront donc le choix des descripteurs du système et permettront de tirer des conclusions selon les valeurs de redondance, fussent-elles fortes ou faibles, l'essentiel étant alors qu'elles soient significatives. On notera que cette condition est nécessairement réalisée puisque chaque prédicteur intervenant dans la procédure divisive est testé selon l'équation 14.

### 5.2 Corrélations entre prédicteurs

L'existence de corrélations entre descripteurs est une réalité incontournable. On verra même plus loin qu'elle est tout à fait souhaitable. Dans l'application décrite plus haut, six variables se rapportent aux températures et il est difficile de penser que les informations qu'elles apportent à la configuration des types de couverts forestiers ne se recoupent pas largement. Par contre, ce qu'on peut affirmer c'est que les résultats de l'analyse (plus précisément la néguentropie et la redondance) sont *exempts de tout effet de corrélation entre prédicteurs*. Cette propriété est dû à la nature combinatoire de la procédure divisive. En effet, si deux variables  $v_k$  et  $v_{k'}$  sont corrélées et si la première  $v_k$

est sélectionnée comme prédicteur, la seconde  $v_{k'}$  ne peut être sélectionnée pour une division ultérieure des sous-ensembles issus de  $v_k$  que si elle apporte une néguentropie additionnelle à celle de  $v_k$ <sup>7</sup>.

Figure 3



La Figure 3 montre que la variable *température d'octobre* fournit la contribution la plus importante à la néguentropie générale (plus de 60%). Les autres températures n'apportent que des contributions mineures (moins de 5%). Si on répète l'analyse en éliminant la variable *température d'octobre* on note alors que la *température annuelle* prend la place la variable écartée avec une contribution proche de 70% à la néguentropie. Ceci montre bien que l'algorithme élimine tout effet de corrélation entre prédicteurs. Ceci montre également qu'il est en fait souhaitable d'introduire des variables partiellement corrélées dans la base de données. En effet, dans cet exemple, les multiples variantes de températures dont les néguentropies se complètent mutuellement, forment une sorte de “*syndrome température*”, accroissant ainsi la néguentropie totale sans toutefois créer de redondance. En utilisant des groupes de variables corrélées, on cernerá donc mieux les relations entre variable-cible et variables descriptives du système.

### 5.3 Désordre et ordre dans le système $U$

- pour accéder à un niveau d'ordre quelconque, un système doit nécessairement posséder une certaine complexité inhérente; à la limite, dans un système uniforme (complexité nulle), la notion d'ordre perd tout sens;
- un système complexe n'est pas nécessairement ordonné; il est cependant potentiellement ordonnable;
- l'ordre dans un système ne peut être appréhendé que par l'intermédiaire d'une relation entre des propriétés distinctes des éléments du système. Il n'y a ordre que si cette relation restreint l'entropie du système et, par conséquent restreint sa capacité de reconfiguration

Revenant au système  $U$ , on sait que l'entropie  $H_{k^*}$  mesure la complexité inhérente à la variable-cible. On sait aussi que la chute d'entropie observée au cours de la démarche de division, c'est-à-dire la néguentropie  $I_{k^*|C}$  et la redondance  $R_{k^*|C}$  associée, mesurent précisément l'ordre introduit par la combinaison de prédicteurs  $C$ . C'est donc l'ordre appréhendé par l'intermédiaire de la relation entre la variable-cible et l'ensemble  $V$  des descripteurs du

<sup>7</sup> Ceci se comprend aisément en considérant l'exemple suivant. Supposons que la base de données contienne deux fois la même variable sous deux noms  $v_k$  et  $v_{k'}$ . Ces deux variables sont donc totalement corrélées. Si  $v_k$  est retenue comme prédicteur, l'ensemble correspondant est divisé selon les classes de  $v_k$  et, par conséquent, les sous-ensembles issus de cette division sont homogènes par rapport aux états de  $v_{k'}$  ce qui entraîne que leur entropie vis-à-vis de  $v_{k'}$  est nulle et *ipso facto* que leur néguentropie est nulle.  $v_{k'}$  ne peut donc intervenir que si elle n'est pas totalement corrélée à  $v_k$  et seulement dans la mesure où elle fournit une néguentropie additionnelle à celle de  $v_k$ . En conclusion, les néguentropies de  $v_k$  et de  $v_{k'}$  ne peuvent redonder en aucune manière.

système. L'appareil mathématique sur lequel repose l'algorithme est donc conforme aux propositions ci-dessus. L'ordre apparaît donc comme la part de complexité rendue prédictible par la connaissance de  $C$  ou, encore, comme négentropie opposée à l'entropie c'est-à-dire la part du potentiel de reconfiguration de la variable-cible restreinte par les contraintes exercées par les descripteurs participant à la combinaison  $C$ .

Tout système tel que  $U$  occupe une position particulière sur une échelle d'organisation qui s'étend entre redondance nulle et redondance complète :

$$0 \leq R_{k^* \setminus C} \leq 1 \quad (17)$$

L'analyse du système apparaît donc à la fois comme une caractérisation des relations entre prédicteurs et variable-cible et des contraintes qui régissent la distribution spatiale de cette dernière et comme une caractérisation du niveau d'organisation (Burel et Baudry, 1999).

## 6 Conclusion

La théorie de l'information et ses antécédents issus de la thermodynamique statistique restent bien le cadre conceptuel privilégié pour comprendre les questions d'organisation en analyse des systèmes (Chaitin, 2001). Mais ce cadre fournit en outre des outils permettant d'aborder des questions sous-jacentes à cette problématique de l'organisation : prédictibilité des états locaux, reconnaissance des contraintes pesant sur la distribution spatiale de ces états et, indirectement, mise en évidence de relations fonctionnelles liant la variable-cible aux autres descripteurs du système.

L'algorithme qu'on vient de voir et son logiciel de mise en œuvre remplissent donc un ensemble de missions qui en font un instrument riche et puissant d'analyse spatiale particulièrement adapté au type de système décrit plus haut. Les conditions actuelles dans lesquelles travaille le chercheur géographe comme la modélisation à base individuelle, conditions rappelées dans l'introduction de cet article, devraient susciter un regain d'intérêt pour les applications de la théorie de l'information dans le domaine de l'espace géographique.

## 7 BIBLIOGRAPHIE

ANTROP M. (1998), Landscape Change - Plan or Chaos, *Landscape and Urban Planning*, vol. 41, n° 3-4, p.155-161.

BAUDRY J. (1989), Interactions Between Agricultural and Ecological Systems at the Landscape Level. *Agriculture, Ecosystems and Environment*, vol. 27, p. 119-130.

BRILLOUIN L. (1951), Physical Entropy and Information II, *Journal of Applied Physics*, vol. 22, n° 3, p. 338-343.

BUREL F., BAUDRY J. (1999) *Écologie du Paysage : Concepts, Méthodes et application*, Paris, Technique et Documentation, p. 166-177.

CHAITIN G. J. (2001), *Exploring Randomness*, London, Springer-Verlag, p. 95-108.

DAVIS F. W., DOZIER J. (1989), Information Analysis of a Spatial Database for Ecological Land Classification. *Photogrammetric Engineering and Remote Sensing*, vol. 56, n° 5, p. 605-613.

DEANGELIS D. L., ROSE K. A. (1992), Which Individual-Based Approach is most Appropriate for a Given Problem? In *Individual-Based Models and Approaches in Ecology, Populations, Communities and Ecosystems*, DEANGELIS D. L., GROSS L. J. eds, New York, Chapman, p. 67-87.

DUMANSKI J., PHIPPS M., HUFFMAN E. (1987), A Study of the Relationships between Soils Survey Data and Agricultural Land Use Using Information Theory, *Canadian Journal of Soil Science*, vol. 67, p. 95-102.

GREY S. L. (1989), Canada's Forest Inventory 1986, Technical Supplement, *Information Report PI-X-86*, Petawawa.

JIANG W. (1996), *A Study of the Relationships Between Forest Distribution and Environmental Variables Using Information Theory : A Regional Scale Model for Predicting Forest Response to Global Warming*. Ph.D. Thesis, University of Ottawa.

PHIPPS M. (1981), Entropy and Community Pattern Analysis. *Journal of Theoretical Biology* . Vol. 93, p. 253-273.

PHIPPS M., BAUDRY J., BUREL F. (1986), Dynamique de l'organisation écologique d'un paysage rural : Modalités de la désorganisation dans une zone péri-urbaine, *Comptes Rendus de l'Académie des Sciences de Paris* , vol.303, n° 7, Ser.III, p. 263-268.

PHIPPS M. (1995), Paysage Écologique, Système Organisé et Analyse de l'Information, dans *Méthodes et Réalisation de l'Écologie du Paysage pour l'Aménagement du Territoire*, Domon G., Falardeau J. éditeurs, Morin Heights, Polyscience Publication Inc., p. 15-24.

PINEDA F.D., DE PABLO C.L., CASADO M.A., DE MIGUEL J.M. (1988), Ecological Structure Recognized by Means of Entropy Analysis : Assessment of Difference between Entropy Values, *Journal of Theoretical Biology* , vol.135. p. 283-293.

WEBBER M. (1977), Pedagogy again : What is Entropy ? *Annals of the Association of American Geographers*, Vol. 67, p. 254-276.

WILSON A. G. (1970), The Use of the Concept of Entropy in System Modelling. *Operational Research Quarterly*, vol. 21, n° 2, p. 247-265.

