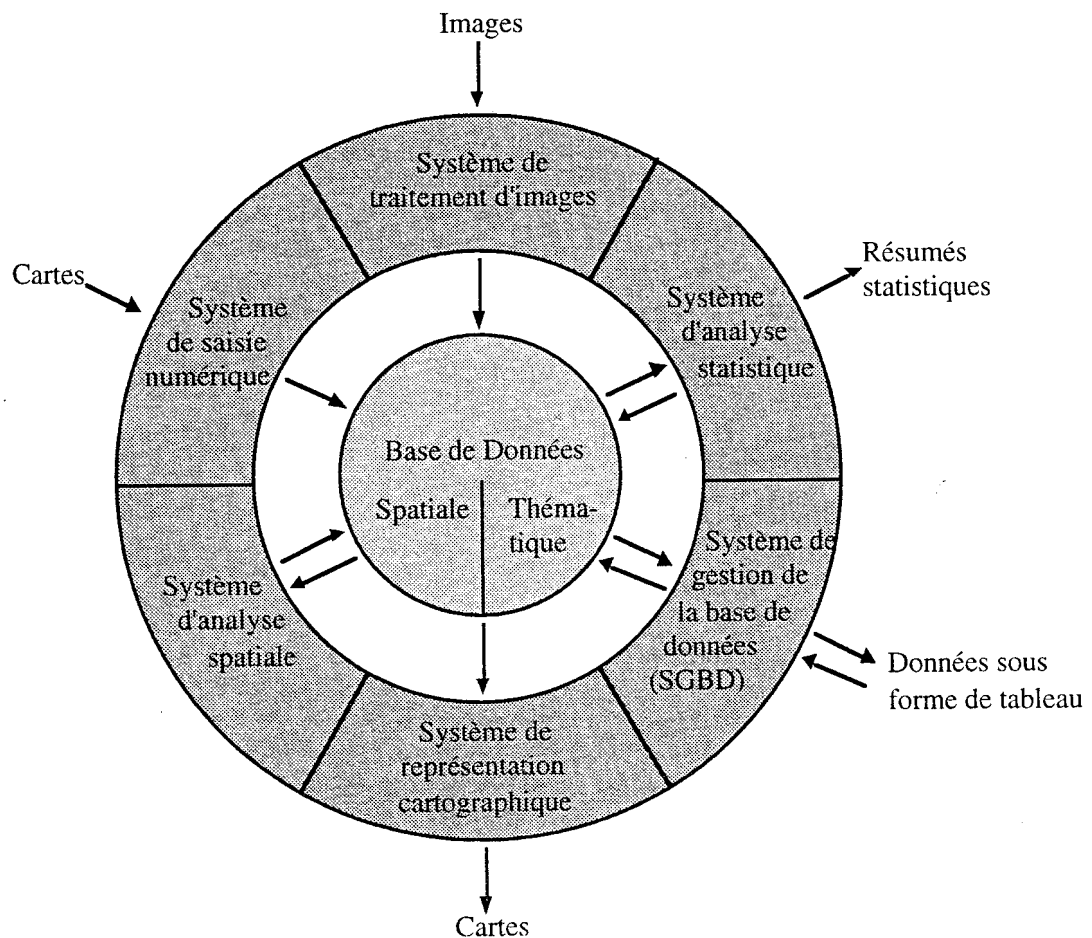


MODELISATION PREDICTIVE

SIG : quelques définitions

Un Système d'Information Géographique (SIG) est un système informatisé d'acquisition, de gestion, d'analyse, et de représentation de données à référence spatiale.

Un SIG intègre plusieurs composantes présentées figure 1 et détaillées ci-dessous.



Les composantes d'un SIG [Eastman, 1993]

□ La base de données à caractère spatial et thématique

Le cœur du système est la base de données géographique (BDG) constituée d'un ensemble numérique de "cartes" et d'informations associées.

Comme cette base de données décrit des objets à la surface de la terre, elle est composée de deux éléments : une base de données décrivant les objets spatiaux (localisation, forme) et une autre définissant les caractéristiques thématiques (attributs) de ces objets.

Par exemple, on peut avoir une base de données décrivant des objets spatiaux tels qu'un ensemble de parcelles, associée à la description thématique de ces dernières telle que l'utilisation du sol, leur propriétaire, leur valeur foncière, etc. qui constitue la base de données thématique.

❑ Système de représentation cartographique

Autour de la BDG gravite un ensemble de systèmes. Le plus élémentaire est celui de représentation cartographique. Ce dernier permet de sélectionner des éléments de la BDG et de les représenter cartographiquement sur de nombreux supports (écran, imprimantes, tables traçantes) et dans différents formats graphiques.

❑ Système de saisie numérique

La saisie numérique d'une information à caractère spatial telle que les documents cartographiques constitue un autre système de base. Celui-ci permet de convertir l'information analogique d'une carte en une information numérique

❑ Système de gestion de la BDG

Traditionnellement, ce terme fait référence à un type de logiciels chargé de gérer des données thématiques. A l'aide d'un SGBD, il est possible d'introduire des informations thématiques, sous la forme de tables ou de statistiques et subséquemment d'en extraire des éléments spécifiques sous les deux mêmes formes. Plus important encore, un SGBD permet l'analyse de ce contenu thématique. En fait, de nombreuses analyses de données spatiales ne comportent pas de vraie composante spatiale, elles se satisfont donc d'un SGBD conventionnel. Le produit final, la carte, sera certes spatial, mais l'analyse elle-même n'a pas de caractéristique spatiale.

❑ Système d'analyse spatiale

L'analyse spatiale de l'information est une extension des capacités d'interrogation des bases de données traditionnelles, en prenant en compte la localisation des observations. En effet l'interrogation traditionnelle d'une base de données fonctionne parfaitement, si les caractéristiques considérées concernent les mêmes entités, sinon l'interrogation est impossible. Pour cette opération, un SIG est nécessaire car il permet de considérer le recouvrement spatial d'entités. Cette procédure, appelée recouvrement ou superposition, correspond à l'opération manuelle de superposition d'une carte transparente et d'une seconde carte.

❑ Système de traitement d'image

En complément aux systèmes de base constituant un SIG, le traitement des images de télédétection inclut des procédures d'analyse statistique qui permet de transformer le contenu originel d'une image de télédétection en une information au contenu thématique exploitable à l'aide de procédures de classification.

❑ Système d'analyse statistique -

Ce système propose à la fois des procédures d'analyse statistique classiques concernant la dimension thématique et des procédures statistiques d'analyse de la dimension spatiale.

Il existe deux techniques fondamentales de représentation numérique des données à caractère spatial: l'approche en *mode objet* (à structure *vecteur*) et l'approche en *mode image* (à structure *raster* ou *en maille*)

Dans une représentation en **mode objet**, les limites ou les bordures des objets sont décrites comme une séquence de points qui, joints par des lignes, forment la représentation graphique de cet objet. Les points eux-mêmes sont décrits par une paire de coordonnées X, Y dans un système de référence. Les attributs ou caractéristiques thématiques de ces objets spatiaux sont eux stockés dans une base de données conventionnelle à l'aide d'un système de gestion de base de données (SGBD).

Par exemple, un ensemble de parcelles décrites en mode objet peut être associée à une base de données thématiques contenant le nom, l'adresse du propriétaire, la valeur foncière et l'affectation. Le lien entre ces deux types d'information - qui constituent deux fichiers de données consiste en un simple identificateur qui est un nombre unique assigné à chaque objet de la carte et que l'on retrouve dans le fichier des attributs (figure ci-dessous).

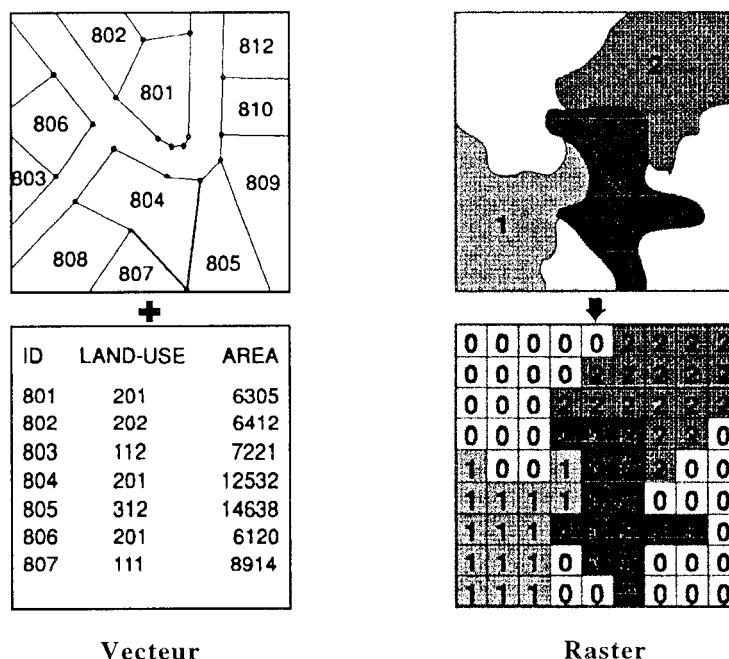
La seconde forme majeure de représentation numérique de l'information spatiale est le **mode image**. Dans un système en mode image, la description des objets et celle de leurs caractéristiques thématiques constituent des fichiers à structure unique.

En fait, les entités spatiales décrites ne sont pas à proprement parlé des objets spatiaux, mais des unités d'observation qui résultent de la subdivision de la zone d'étude en un maillage de cellules rectangulaires (figure ci-dessous).

Chaque cellule contient une valeur numérique qui peut représenter soit un identificateur d'appartenance à un objet spatial, soit une catégorie, classe ou valeur thématique.

Par exemple, une cellule affectée de la valeur "6" peut indiquer soit qu'elle appartient au district n°6, soit que le type de sol qui la constitue appartient à la catégorie 6, soit encore que son altitude se situe à 6m au-dessus du niveau de la mer.

Bien que l'information stockée dans cette structure en maille (*raster*) ne soit pas nécessairement visible dans le paysage, elle est appelée *image* et les éléments qui la constituent, les cellules, sont aussi appelés *pixels*. Pour visualiser ces images à l'écran ou sur support papier, les valeurs de leurs pixels sont simplement remplacées par des symboles graphiques tels que des couleurs, des tons de gris ou des textures.



La représentation numérique des données [Eastman, 1993]

Ces deux systèmes ont des points forts inhérents à leur manière de représenter et de gérer l'information spatiale. Les systèmes d'approche image gèrent de grandes quantités d'informations malgré des techniques efficaces de compression, car ils stockent l'information de chaque cellule d'une image, quel que soit son intérêt.

Par contre, l'avantage est que la zone d'étude est découpée en unités régulières et uniformes, qui rendent ces systèmes plus performants pour des tâches d'analyse ; particulièrement lorsque les distributions sont spatialement continues (MNT).

Le second avantage de la structure raster est sa similitude avec l'architecture interne des ordinateurs ; ainsi l'évaluation de problèmes faisant appel à la combinaison mathématique de plusieurs grilles (images) est rapide. Cette approche est efficace pour l'application de modèles numériques environnementaux. De plus, comme les données de télédétection ont cette structure raster, elles peuvent être directement traitées par un tel système.

Les systèmes en mode objet offrent une plus grande efficacité dans la gestion de cette information spatiale. Ces derniers stockent l'information de manière compacte car ils ne retiennent que la description des contours des objets.

De plus, la représentation cartographique des objets étant directement liée à leurs contenus thématiques stockés dans la base de données, les systèmes vecteur permettent d'interroger chaque objet spatial de manière interactive à l'écran en pointant dessus pour connaître ses caractéristiques spatiales et thématiques. Ils produisent des cartes thématiques simples à partir d'interrogations. En comparaison à l'alternative image, l'approche objet n'offre pas de grandes capacités d'analyse de distributions spatialement continues.

Le logiciel IDRISI utilisé intègre des éléments de ces deux modes de représentation numérique. Bien qu'il soit avant tout un système raster, IDRISI utilise la structure vecteur comme forme principale d'importation des données.

Apports des SIG en archéologie spatiale

Les deux thématiques principales auxquelles s'intéressent les archéologues, à savoir les paramètres environnementaux et l'exploitation agricole du territoire, sont longtemps restées distinctes et indépendantes. L'ampleur de la tâche et la multiplicité des paramètres entrant en ligne de compte produisaient malgré un effort certain, une segmentation des études et des résultats.

Le SIG est un outil de diffusion récent, qui ouvre de nouvelles voies à la recherche en archéologie spatiale. La souplesse et la puissance du support informatique permettent de stocker et de combiner des données d'origines diverses (cartes, images satellites, prospections...) et de nature variée (données environnementales, historiques, archéologiques).

La confrontation et la combinaison de ces données multiples offrent des opportunités nouvelles en matière d'analyse spatiale.

Insérées au sein d'un SIG, ces études complémentaires permettent la spatialisation et la modélisation des dynamiques anthropiques et environnementales. Au terme de cette opération, l'organisation et l'exploitation du milieu pourront être perçues et synthétisées sur une échelle diachronique.

Cette technique offre deux avantages fondamentaux pour les applications archéologiques:

- Le traitement d'un nombre très élevé de données (trop long et lourd pour traiter à la main), notamment intégrant des paramètres archéologiques et environnementaux.
- La prise en compte globale du territoire (et pas seulement du site) comme une entité archéologique, de manière à pouvoir juger du poids réel des paramètres environnementaux et sociaux dans la localisation des sites.

Les études dans le domaine de l'environnement ont montré l'importance d'une approche multidisciplinaire ainsi que d'une prise en compte de la dimension spatiale.

Le caractère multidisciplinaire intervient à la fois au niveau de l'information collectée sur la zone d'études et à celui de l'analyse. La dimension spatiale quant à elle, rend compte des variations des phénomènes dans la surface de cette zone, de leurs interactions, ainsi que de l'arrangement spatial des objets qui la composent.

Problématique et méthodologie

Après des recherches bibliographiques menées aussi bien sur des documents et rapports archéologiques de travaux similaires que sur des ouvrages de lois statistiques, la première étape du travail a été de déterminer la démarche à employer pour réaliser un modèle prédictif.

La démarche de cette étude est orientée vers la gestion du patrimoine et non vers une approche archéologique d'étude de peuplement. En effet, les résultats des tests statistiques seront analysés et associés, lorsqu'ils peuvent l'être, à des faits établis dans des ouvrages et

articles archéologiques d'étude des peuplements; mais le modèle prédictif se limitera à la détermination de zones où la probabilité de découverte de gisements archéologiques est forte et des zones où cette probabilité est faible.

Le test statistique employé est le test du χ^2 de Pearson qui est aussi le plus employé en archéologie prédictive pour sa simplicité et ses bons résultats empiriques. De plus, celui-ci est bien adapté dans le cadre de notre étude car l'effectif de notre base de données est suffisant.

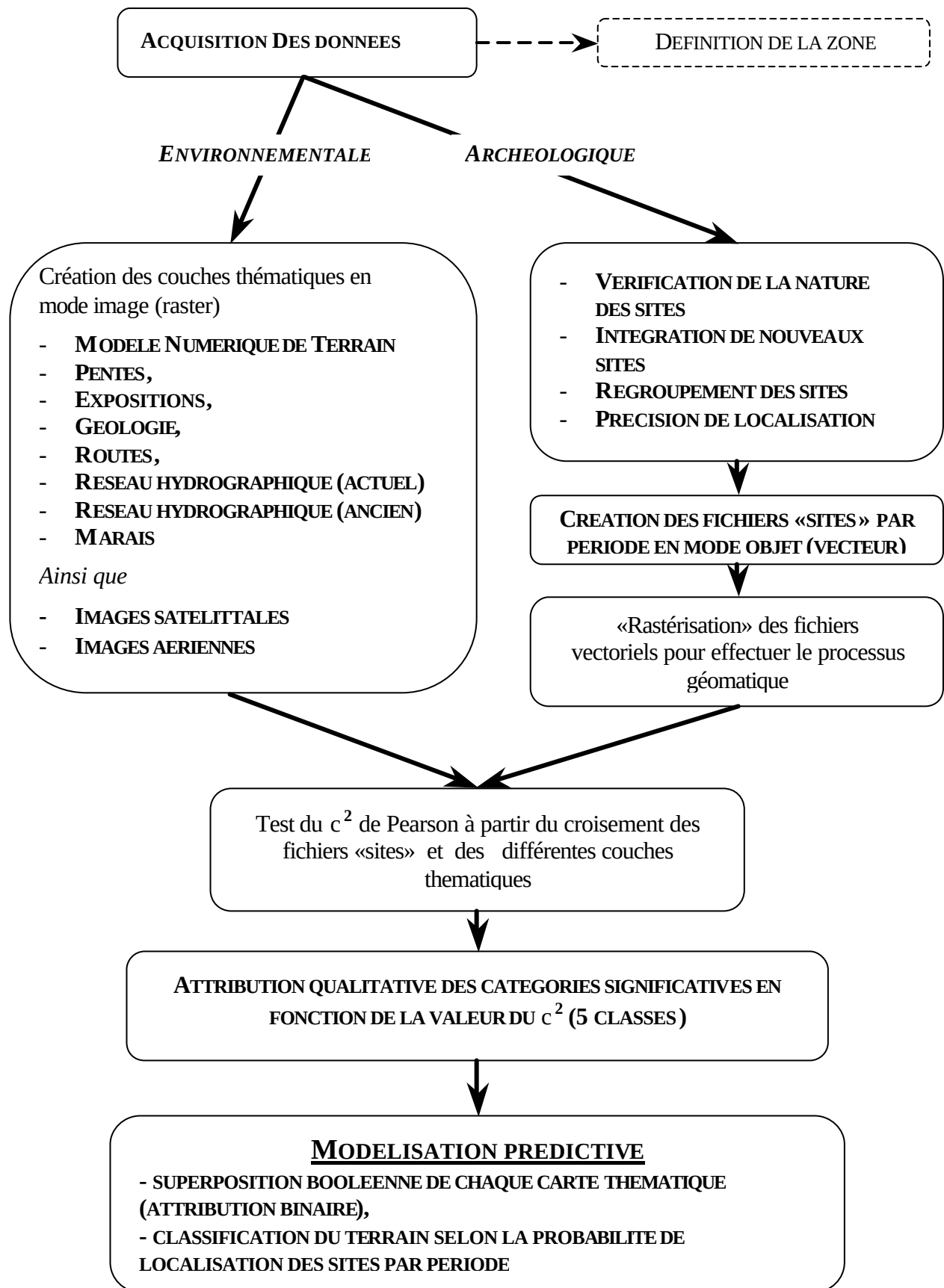
Une fois le cadre du projet déterminé, il faut mettre en place les moyens techniques pour réaliser le projet précédemment décrit. Si les logiciels de SIG constituent un mode de traitement puissant et moderne, il ne s'agit pas d'une panacée universelle et leur usage requiert l'emploi d'une méthodologie stricte. La problématique mais également, la sélection, la saisie et le traitement des informations sont autant de choix préalables qui conditionnent l'obtention de résultats productifs et scientifiquement exploitables.

Face à la multitude et à la diversité des techniques offertes par les logiciels existants, il est impératif de connaître leur fonctionnement et leurs conditions d'application pour que leur emploi débouche sur des résultats cohérents. Ainsi, pour réaliser chacun des processus de sélection, saisie et traitement des données, nous avons eu recours à des logiciels aussi divers que :

- IDRISI32 release32
- Arcview 3.2
- Adobe Illustrator 10
- Pathfinder office 2.91
- Microsoft Excel
-

L'élaboration d'un modèle prédictif est un travail pluridisciplinaire faisant appel à tout les aspects d'un SIG.

Chaque opération nécessite une réflexion préalable et une préparation minutieuse : les procédures ont donc été envisagées sous un aspect théorique puis pratique, afin de mieux cerner les impératifs de chacune (voir organigramme ci-dessous).



Statistique : test du χ^2

Notre modèle prédictif se fonde sur l'inventaire des sites archéologiques issus des travaux de prospection menés dans notre région d'étude, et sur les caractéristiques environnementales du territoire.

En présence d'une distribution de sites que nous considérerons à priori aléatoire, nous voulons tester l'hypothèse de sa normalité.

Nous cherchons alors quelle est la distribution théorique que nous aurions du observer dans le cas d'une loi normale de moyenne m et d'écart-type s .

L'effectif théorique (nt) est proportionnel à la surface occupée dans la zone d'étude par l'élément considéré. Les nt sont calculés de la manière suivante:

$$nt = \frac{n_{\text{total}} * s}{s_{\text{total}}}$$

où:

- n_{total} représente le nombre total de sites sur la zone étudiée
- s représente la surface de la catégorie concernée
- s_{total} représente la surface totale de la zone étudiée

Nous pouvons alors comparer :

- les nk observés
- les nt théoriques

Si la distribution observée est normale, il ne doit y avoir aucun écarts entre les nk et les nt . Comme cette condition est rarement remplie, nous étudions alors la variable constituée par ces écarts:

$$c^2 = \sum_k \frac{(nt - nk)^2}{nt}$$

Les lois de c^2 diffèrent légèrement en fonction du nombre de classes k . Si la distribution est normale, plus les classes sont nombreuses, plus les écarts peuvent être grands puisque les occasions d'écart sont multipliées. Les tables de la lois de c^2 sont donc établies en fonction du nombre de classe ou plus exactement de degrés de liberté.

En effet dans le cas d'une loi normale, nous travaillons sur le même effectif total (N), la même moyenne (m) et le même écart type (s) dont la détermination n'est plus libre puisque donnée par avance. Ces trois paramètres qui définissent notre relation sont donc liés et restreignent la « liberté » de notre distribution.

Si notre variable contient k classes, le degré de liberté « dl » de notre distribution par rapport à cette variable sera :

$$dl = k - 3$$

La table du χ^2 fournie en annexe indique si la tendance est significative.

Les résultats empiriques sont les suivants :

Couches environnementales	Nombre de catégories	degrés de liberté (dl)	seuil à 1% (influence significative au delà)
L'occupation du sol	5	2	$\chi^2 = 9.21$
La géologie			
L'altitude	7	4	$\chi^2 = 13.28$
La pente	8	5	$\chi^2 = 15.09$
La distance pondérée au plus proche marais			
La distance pondérée au réseau hydrographique ancien			
La distance pondérée au réseau viaire			
L'exposition	9	6	$\chi^2 = 16.81$

La notion de seuil souligne le fait qu'un test statistique n'est jamais une certitude mais un choix. En effet, choisir un seuil de 1% implique que nous avons tout de même 1 chance sur 100 pour qu'un tel résultat soit dû uniquement aux aléas de l'échantillonnage, mais que nous choisissons de tout de même de considérer notre variable comme **significative**.

Si la variable a une influence significative sur la distribution, on peut alors attribuer le terme « favorable » aux catégories où **nk observés** > **nt théoriques** et « défavorable » si **nk observés** < **nt théoriques**

L'utilisation des tests statistiques en archéologie a donné lieu à de nombreuses polémiques, certains argumentant que les sciences humaines ne pouvaient être régies par de simples formules mathématiques. Cependant des études empiriques ont montré que le test du χ^2 est, dans de nombreux cas, concluant. C'est pour cela qu'il est le plus utilisé en archéologie, même par les archéologues les plus récalcitrants à l'usage de la statistique.

Cependant, l'utilisation du test du χ^2 présente des limites de validité :

- ❑ la proportion de sites présents sur l'unité de terrain considérée doit être supérieure à la proportion de superficie totale que représente cette unité (souvent non vérifié)
- ❑ le nombre total des sites doit être supérieur ou égal à 40
- ❑ si le nombre de classes >2, moins de 20% des classes doivent avoir un effectif théorique <5

Sur la base de ce traitement, on peut établir une classification favorable/défavorable en terme de probabilité de localisation ou de non-localisation de sites pour chacune des catégories des variables environnementales significatives.

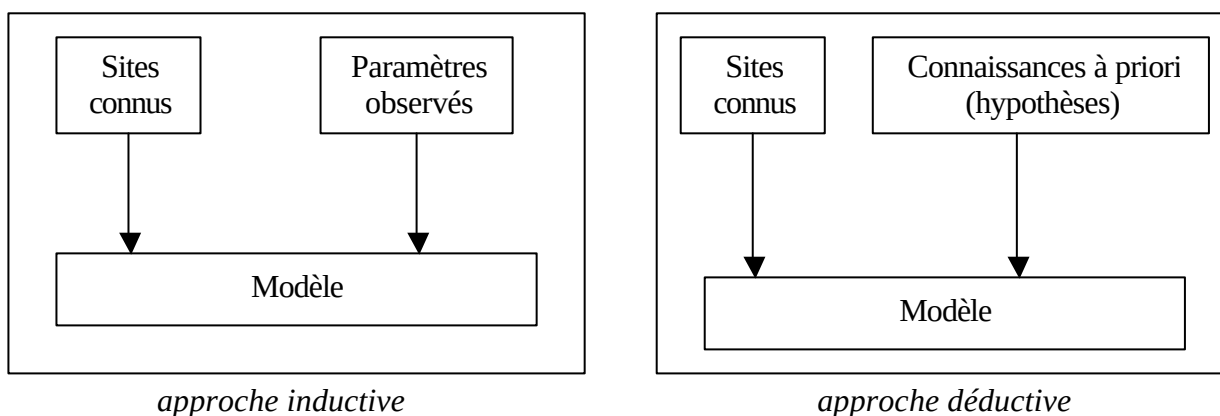
Modélisation prédictive de localisation des sites archéologiques

Le terme de "modèle prédictif " a été créé par les agences d'aménagement du territoire aux Etats-Unis entre 1979 et 1983. Il permet d'évaluer l'impact dans le "paysage culturel" de l'urbanisation, l'extension des réseaux de transport et autres agents destructifs; ou encore de gérer les travaux de prospection (optimisation du temps et des budgets).

Il permet aussi de mener des études scientifiques d'analyse de l'archéologie spatiale: déterminer les facteurs de localisation des sites archéologiques dans le temps et intégrer ces facteurs dans l'étude archéologique.

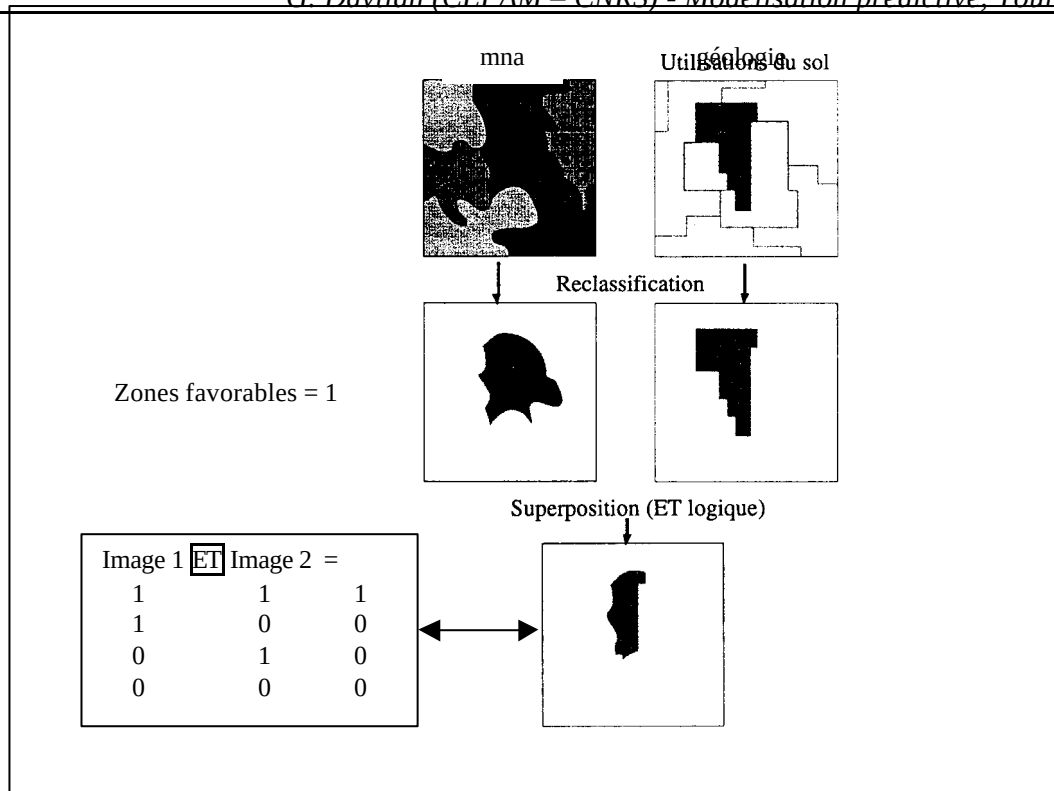
Le modèle prédictif repose techniquement sur des opérations d'analyse statistique et spatiale effectuées sur des images. La structure raster est particulièrement bien adaptée à ce type d'opérations car la description des objets et celle de leurs caractéristiques thématiques constituent des fichiers à structure unique. Il est donc nécessaire de convertir tous les fichiers intervenant dans la modélisation au format raster.

Les prédictions peuvent être faite à partir de paramètres environnementaux observées et de sites connus, c'est le cas d'une approche **inductive** ou à partir d'hypothèses sur le comportement humain, c'est le cas d'une approche **déductive**.



Modélisation inductive

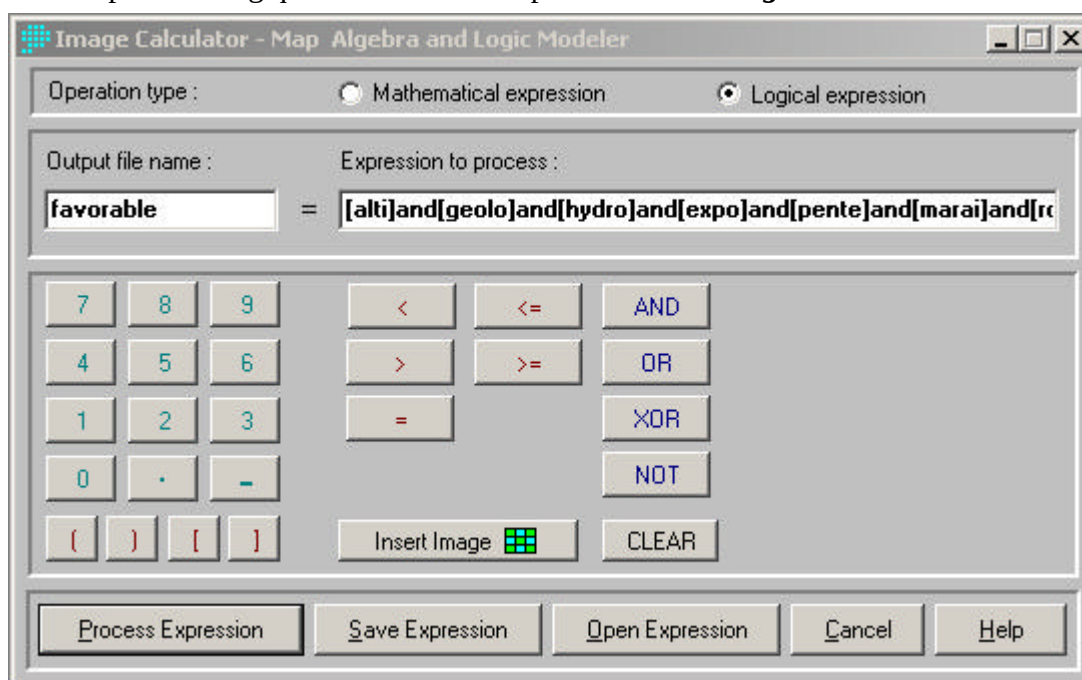
La modélisation inductive consiste en une superposition booléenne de toutes les couches environnementales significatives avec l'opérateur logique ET qui fonctionne de la manière suivante



Opérateur logique ET [Eastman, 1993 modifié]

L'attribut « 1 » est attribué uniquement aux pixels de l'image résultante si tous les pixels homologues sur toutes les images de données environnementales ont pour valeur « 1 ».

Les opérations logiques sont accessibles par la fonction **image calculator** d'IDRISI

Opérations logiques du module **image calculator** d'IDRISI

Les opérations logiques fonctionnent uniquement avec des images booléennes, nous allons donc travailler en 3 temps :

- ❑ Création d'une carte des zones à **forte probabilité** de localisation de sites :

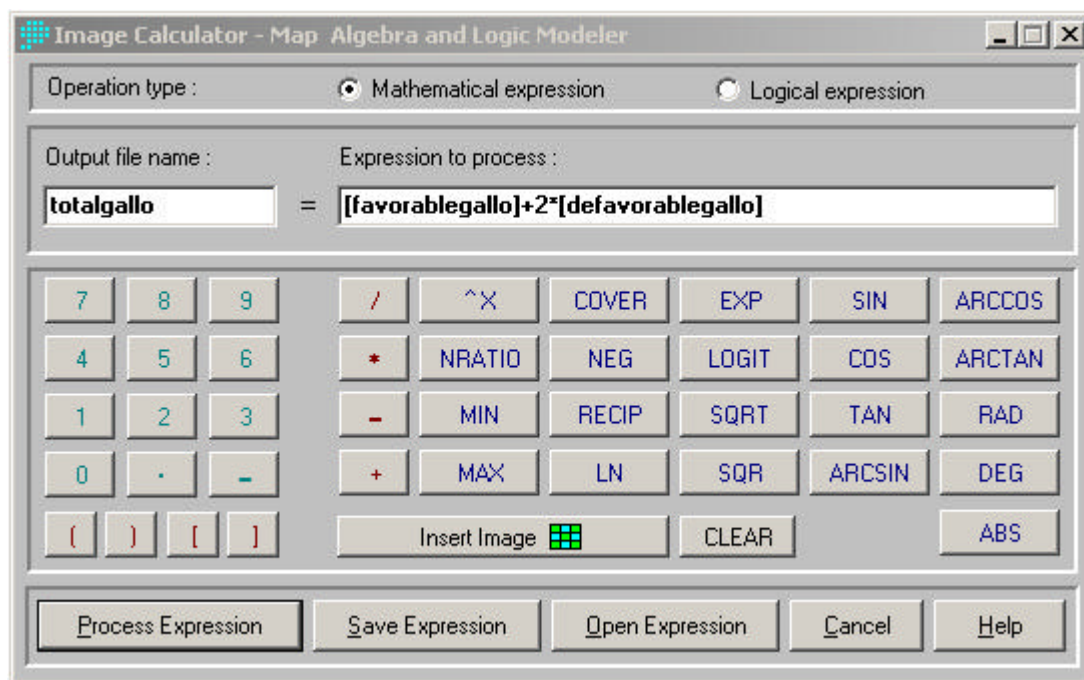
Toutes les couches environnementales sont reclassées en assignant une valeur de «1 » aux zones à forte probabilité de localisation de sites et une valeur «0 » à tout le reste. On superpose ensuite toutes ces couches en utilisant l'opérateur logique **ET** ; on obtient de cette façon la carte des zones à forte probabilité de localisation de sites que l'on nommera « **Favorable** »

- ❑ Création d'une carte des zones à **faible probabilité** de localisation de sites :

De la même façon, toutes les couches environnementales sont reclassées en assignant une valeur de «1 » aux zones à faible probabilité de localisation de sites et une valeur «0 » à tout le reste. On superpose ensuite toutes ces couches en utilisant l'opérateur logique **ET** et on obtient la carte des zones à faible probabilité de localisation de sites que l'on nommera « **Défavorable** »

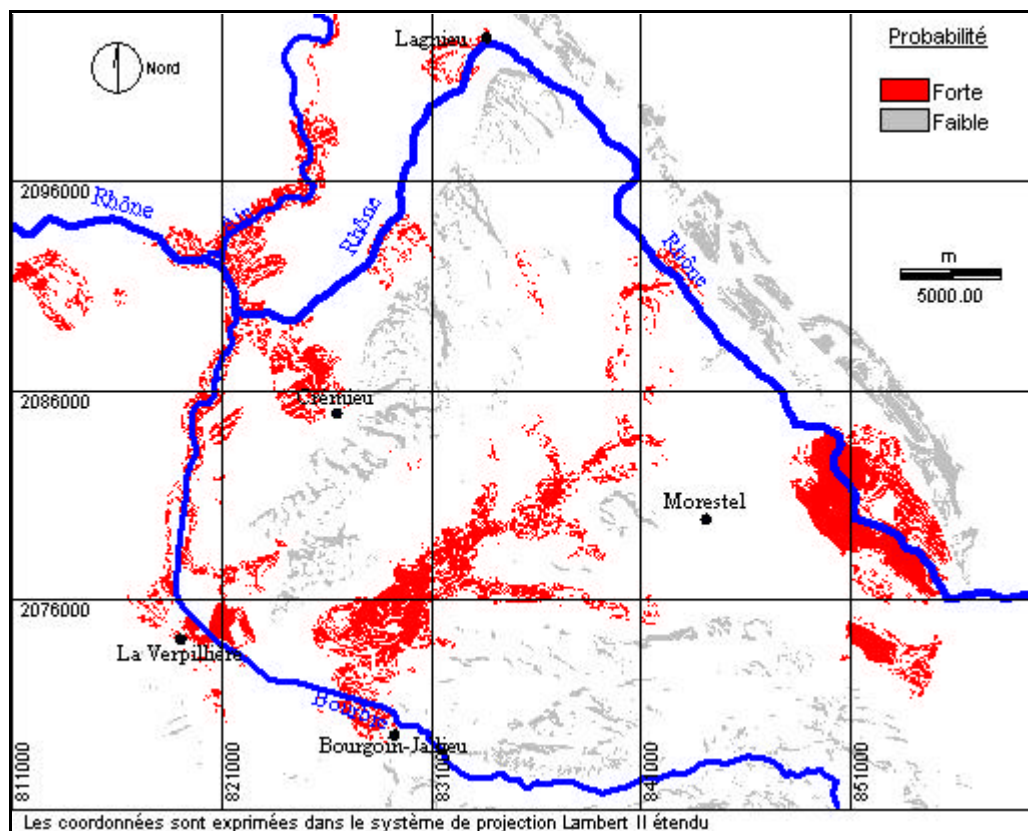
- ❑ Réunion des deux cartes en une seule :

On ajoute alors les deux images «Favorables » et «Défavorables » à l'aide de l'opérateur mathématique « + ». Pour pouvoir identifier ces deux catégories sur l'image résultante on multiplie préalablement l'image « Défavorable » par 2.

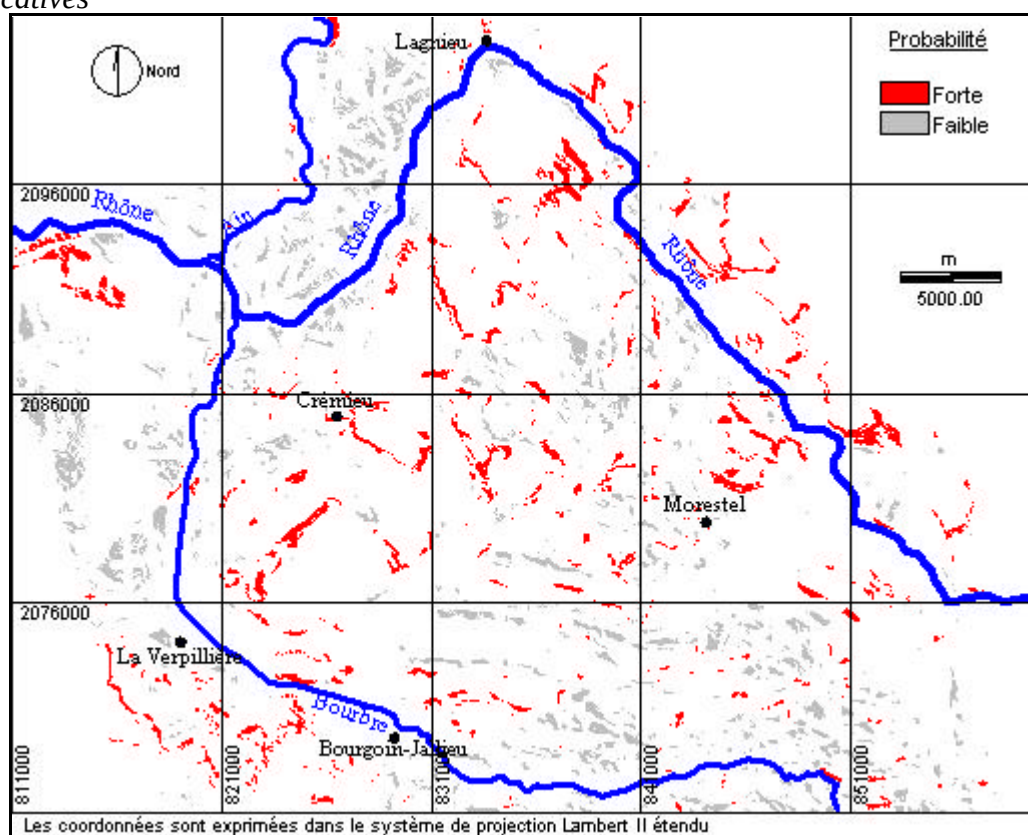


Opérations mathématiques du module **image calculator** d'IDRISI

On obtient ainsi des modèles prédictifs obtenus par superposition de toutes les couches environnementales de manière équivalente.



Carte des probabilités de localisation de sites gallo-romains obtenue par approche inductive: superposition booléenne des 7 couches de données environnementales significatives



Carte des probabilités de localisation de sites médiévaux obtenue par approche inductive: superposition booléenne des 8 couches de données environnementales significatives

Modélisation déductive

Les résultats obtenus jusqu'ici donnent à penser que les contraintes imposées par l'environnement ont une influence primordiale sur l'implantation des sites. S'il s'agit là d'une affirmation exacte pour la période gallo-romaine, il faut toutefois la nuancer, et reconnaître le rôle de l'organisation féodale du territoire et des contraintes climatiques sur l'occupation humaine du milieu pour les époques médiévales et modernes.

Ainsi nous avons pu constater que l'implantation des sites est fortement liée aux contraintes naturelles à l'époque gallo-romaine. En revanche pour la période médiévale des contraintes féodales ont pour conséquence de favoriser certains facteurs (pente, altitude), lors du choix de l'implantation des sites.

De plus l'implantation des sites à la période moderne est tributaire de contraintes climatiques qui mettent au premier plan des facteurs environnementaux, telles l'altitude et la pente ; au détriment de l'exposition et de la géologie. Certaines variables restent cependant invariablement très importantes dans le choix d'implantation des sites tel les réseaux hydrographiques et les réseaux routiers.

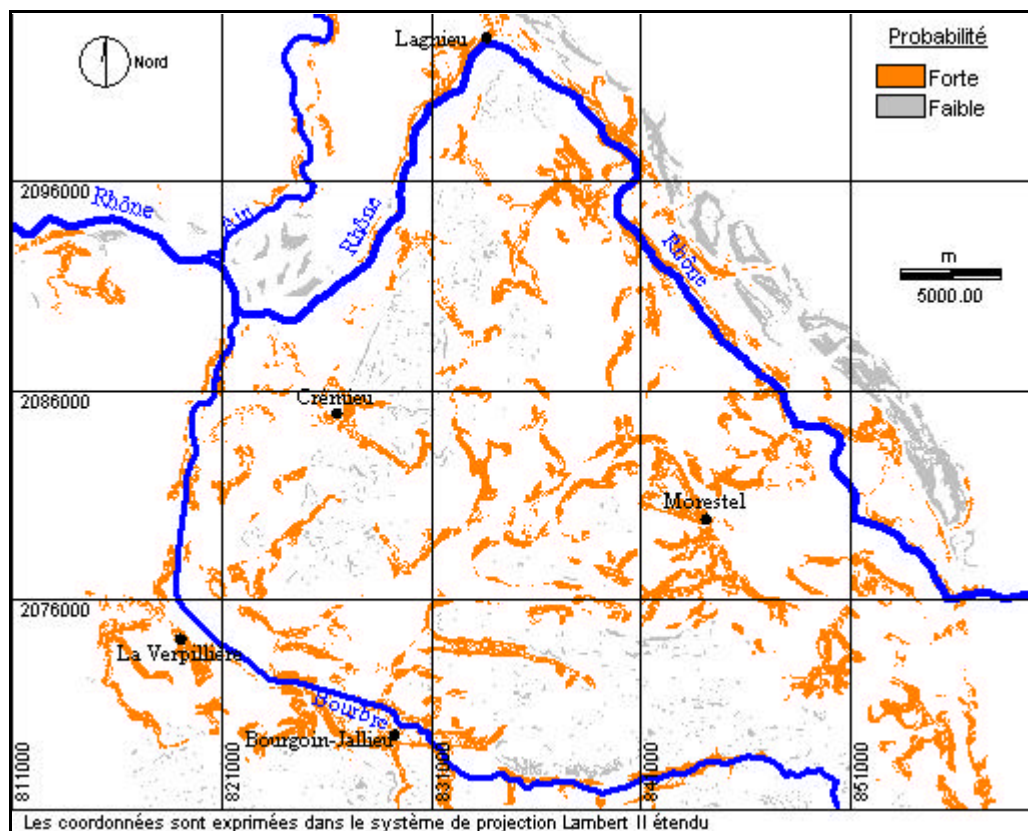
Nous avons donc dressé un tableau s'appuyant sur les analyses statistiques de la répartition des sites et intégrant des hypothèses. Nous attribuerons ainsi un « poids » aux différentes couches environnementales en fonction de leur impact sur le modèle final.

Nous obtenons ainsi les résultats suivants.

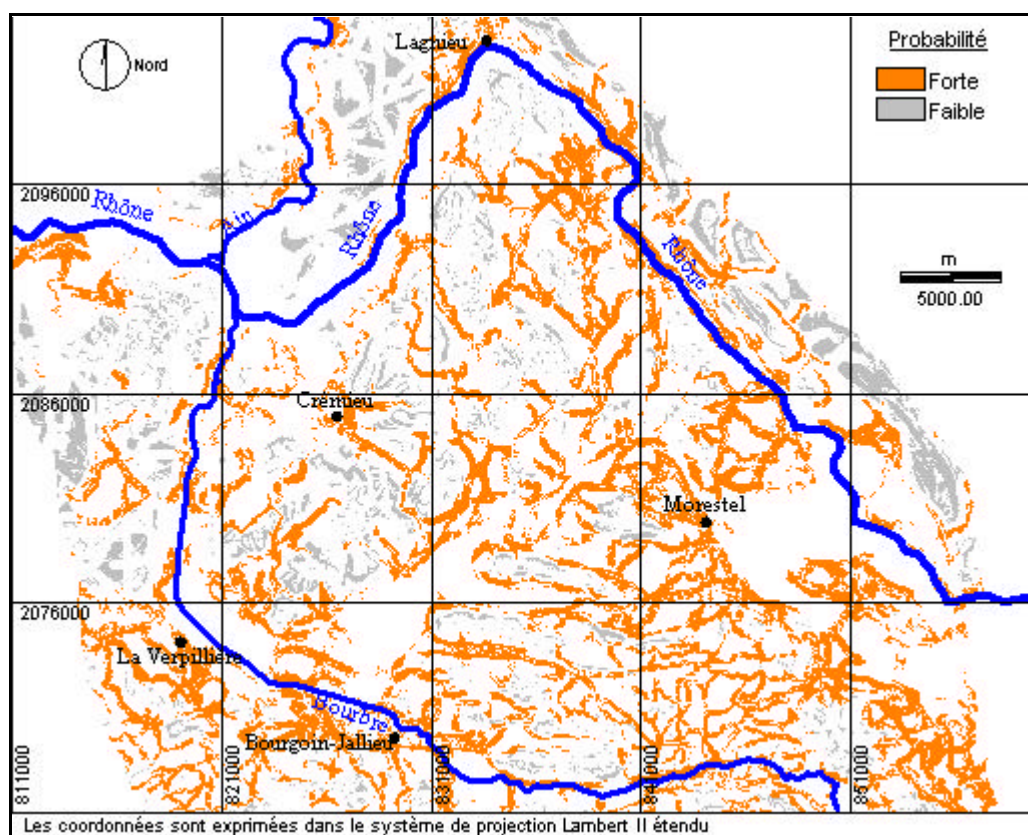
Variables	Pondération		
	Gallo-romain	Médiéval	Moderne
La géologie	2	1	1
l'exposition	2	1	1
L'altitude	2	2	2
la pente	2	2	2
La distance pondérée aux marais	2	1	0
La distance pondérée au réseau hydrographique	2	2	2
La distance pondérée au réseau viaire	0	2	2

A la vue des coefficients affectés à chaque variable, nous remarquons que le modèle existant semble bien adapté à l'époque gallo-romaine, mais pas aux autres époques.

On effectue alors d'autres superpositions booléennes avec uniquement les couches environnementales les plus importantes ; c'est à dire pour la période médiévale : les couches [pente], [route], [hydro], [MNA] et pour la période moderne les couches [pente], [route], [hydro], [MNA].



Carte des probabilités obtenue par approche « déductive » pour l'époque médiévale : superposition booléenne des couches [pente],[route],[hydro],[MNA]



Carte des probabilités obtenue par approche « déductive » pour l'époque moderne : superposition booléenne des couches [pente],[route],[hydro],[MNA]

Modèle inductif ou déductif ?

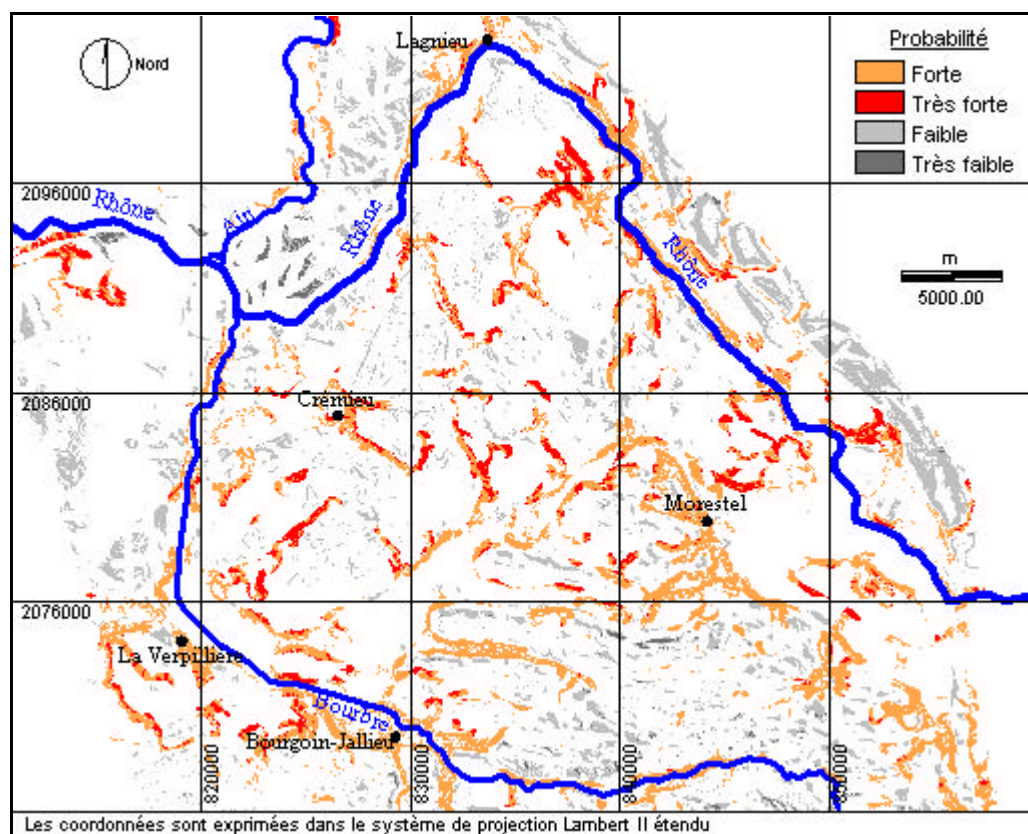
La **modélisation inductive** s'appuie une démarche logique, quasi-automatique et sur des fondements solides de données observées. Mais il est quelquefois difficile d'obtenir des résultats cohérents lorsque l'échantillonnage des sites est non-aléatoire et les données environnementales mal connues. La **modélisation déductive** par contre, n'a pas un cadre bien formalisé et les modèles créés par cette démarche sont plus subjectifs.

Dans la pratique, beaucoup de modèles sont des « hybrides ».

En intégrant le modèle déductif au modèle inductif, nous avons alors créé un modèle avec des zones de **probabilité très forte** obtenues par démarche inductive, de **probabilité forte** obtenues par démarche déductive ; et inversement, des zones de **probabilité très faible** et des zones de **probabilité faible**.

Les modèles obtenus à l'aide des deux démarches allient la rigueur de l'approche inductive, tout en remédiant à certaines incohérences en intégrant des hypothèses.

En observant les cartes prédictives, on remarque que les principales zones urbanisées sont des zones à forte ou très forte probabilité de localisation de sites aux époques médiévale et moderne. Ceci semble bien refléter la réalité car l'habitat actuel est ancré sur les mêmes sites depuis la fin de l'Antiquité. De plus la désertification des basses plaines alluviales survenue à la fin de l'Antiquité est bien visible. Ces modèles «hybrides» semblent donc bien représenter les tendances énoncées lors des analyses statistiques.



Carte des probabilités de localisation de sites médiévaux obtenue par approche « hybride »

Quand utiliser les cartes prédictives ?:

- Pour influencer les cartes d'occupation du sol (POS)
- Pour estimer les coûts d'intervention archéologique
- Pour orienter les choix méthodologiques et techniques, principalement lors des phases de prospection.

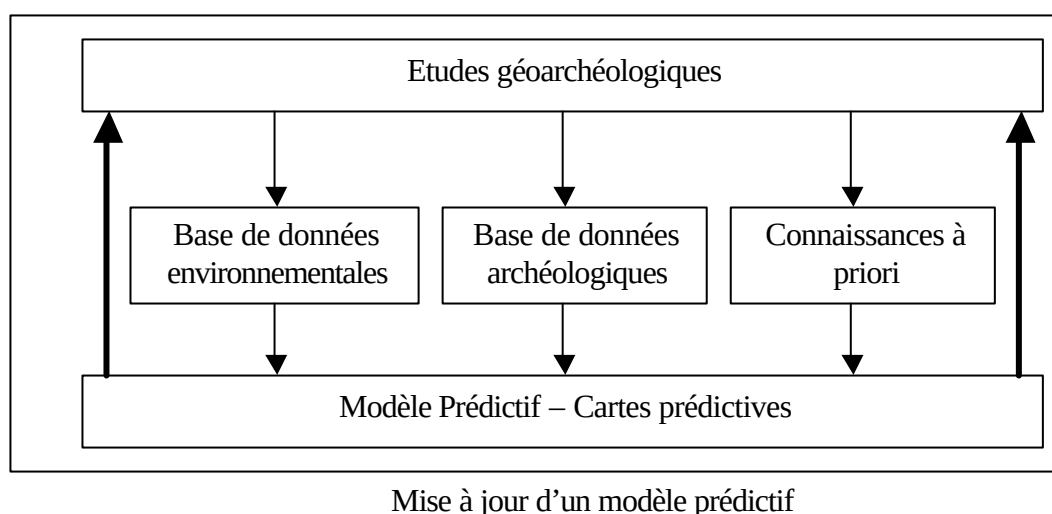
Utilité et limites des modèles prédictifs

Les estimations prédictives sont à considérer d'un point de vue **scientifique** : amélioration des connaissances sur l'évolution de l'occupation des sols; mais aussi **patrimonial** : prévision du potentiel archéologique enfoui dans les basses plaines soumises à d'incessants travaux d'aménagement du territoire.

Elles intéressent à ce titre les laboratoires du CNRS, mais aussi ceux du ministère de la culture, des collectivités locales, et de l'INRAP (Institut National de Recherches Archéologiques Prédictives) qui ont en charge la gestion des différentes cartes archéologiques nationales, départementales, urbaines et les opérations d'archéologie préventive.

Des collaborations sont ainsi à envisager dans les prochaines années, à l'échelon national entre ces organisations pour optimiser la gestion des cartes archéologiques régionales pour développer des modèles de risque archéologique majeur sous SIG, permettant de répondre aux attentes des aménageurs et faciliter ainsi la politique d'aménagement du territoire.

Mais les orientations et décisions reposant sur des cartes prédictives sont limitées, si elles ne reposent pas en amont sur de solides études géoarchéologiques. De plus, les modèles prédictifs se doivent d'être en perpétuelle évolution et d'intégrer chaque nouvelle découverte dans leur base de données. Il ne suffit donc pas de créer un modèle à une date donnée mais aussi de se donner les moyens de le mettre à jour régulièrement.



Avenir et perspectives de développement des modèles prédictifs

La maîtrise de la complexité croissante des problèmes sur le territoire nécessite de disposer d'un ensemble de méthodes et d'outils puissants et adéquats. Ces moyens doivent non seulement traiter rapidement les données à référence spatiale d'origine et de natures diverses, mais aussi les analyser afin d'obtenir une information pertinente, permettant d'éclairer les décideurs dans leur prise de décision. Parmi les solutions offertes, plusieurs organismes gouvernementaux privés concernés par le territoire se sont appuyés sur les systèmes d'information géographiques. Cependant, malgré son importance, le type d'information fourni par les SIG actuels s'est avéré insuffisant et parfois inadéquat pour éclairer le décideur particulièrement au niveau stratégique. De nombreuses recherches ont porté sur la proposition de nouvelles solutions conceptuelles et méthodologiques appelée « **analyse multicritère** ».

Mais ces solutions se sont avérées limitées quant à leur apport réel aux décideurs car elles se heurtent aux limites des modèles de données sur lesquels sont bâtis les SIG et à la logique booléenne qui régit leurs opérations. De plus, l'information rendue disponible par la base de données à référence spatiale n'est pas suffisante pour rendre compte de tous les besoins en informations nécessaires pour le processus décisionnel : les informations manquantes sont généralement d'ordre politique, socio-économique et plus souvent d'ordre heuristique (l'expérience du décideur est importante), soit des facteurs essentiels pour la réussite du processus décisionnel.

Or, l'information, qui est fournie actuellement par les SIG, se limite d'une part à des contraintes qui représentent les ressources disponibles et connues, et, d'autre part, à des attributs décrivant les variantes de choix. Ces informations sont de nature opérationnelle et diffèrent des informations stratégiques qui sont plutôt « préférentielles ».

Les efforts des chercheurs dans le domaine de l'analyse multicritère se sont portés sur le manque de fonctionnalités analytiques des SIG. En effet, l'analyse spatiale offerte par les SIG actuels n'incorpore pas tout le potentiel nécessaire et souhaitable pour supporter le processus décisionnel à référence spatiale.

De plus, les domaines d'application d'un SIG sont si variés qu'il semble, à ce jour, difficile d'automatiser des choix propres à une étude bien précise par le développement d'un outil générique. Cependant, des travaux de recherche pourraient améliorer la correspondance entre les caractéristiques des problèmes à référence spatiale et les conditions d'utilisation des méthodes multicritères. Mais le problème ne réside pas seulement dans les limites de l'analyse spatiale dans les SIG, mais aussi dans la manière dont les SIG influencent la modélisation et l'appréhension des données à référence spatiale et dans les limitations inhérentes à la base de données à référence spatiale elle-même. On pourrait alors essayer, par exemple, de modéliser de façon plus adéquate les données disponibles comme les données environnementales ou archéologiques, pour parvenir à mieux les intégrer dans la Base de Données à Références Spatiale.

Enfin, nous espérons que ce mémoire participera à rapprocher les communautés de l'archéologie et des SIG ou, du moins, à susciter de l'intérêt pour la modélisation prédictive qui lie sciences humaines et techniques.