

1989
ID
16

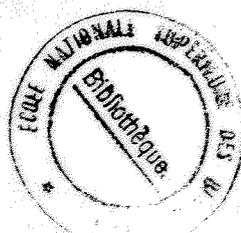
1

**NOTE DE SYNTHÈSE PRÉSENTÉE POUR L'OBTENTION DU
DESS D'INFORMATIQUE DOCUMENTAIRE
À L'UNIVERSITÉ CLAUDE BERNARD LYON I**

par

VIEVILLE-JURBERT Jocelyne

**CARTOGRAPHIE DE LA VULNERABILITÉ DES EAUX
SOUTERRAINES AUX POLLUANTS ET CONCEPTION ASSISTÉE PAR
ORDINATEUR**



6 Juillet 1989

Directeur de mémoire : Mme SUAIS M.F., BRGM LYON
Directeur universitaire : Pr BOUCHE R.

1989
ID
16

**NOTE DE SYNTHÈSE PRÉSENTÉE POUR L'OBTENTION DU
DESS D'INFORMATIQUE DOCUMENTAIRE
À L'UNIVERSITÉ CLAUDE BERNARD LYON I**

par

VIEVILLE-JURBERT Jocelyne

**CARTOGRAPHIE DE LA VULNERABILITÉ DES EAUX
SOUTERRAINES AUX POLLUANTS ET CONCEPTION ASSISTÉE PAR
ORDINATEUR**



6 Juillet 1989

Directeur de mémoire : Mme SUAIS M.F., BRGM LYON
Directeur universitaire : Pr BOUCHE R.

1989

ID
16

SOMMAIRE

	<u>pages</u>
INTRODUCTION.....	4
1. LE CHOIX DES DONNEES PARTICIPANT A L'ELABORATION DES CARTES DE VULNERABILITE.....	5
1.1. Les contraintes concernant l'aspect spatial des données.....	5
1.2. Les données descriptives en relation avec le thème de vulnérabilité des nappes.....	5
1.2.1. Les données concernant le site.....	5
1.2.2. Les données concernant les polluants.....	6
1.2.3. Les données concernant l'utilisation du site et des alentours.....	8
2. LA SAISIE DES DONNEES.....	9
2.1. La saisie des données graphiques.....	9
2.1.1. La saisie sur table à numériser.....	9
2.1.2. La saisie par balayage.....	9
2.1.3. La saisie par suivi automatique du trait...	9
2.1.4. Les corrections et modifications compagnant la saisie.....	10
2.2. La saisie des données descriptives.....	10
3. STOCKAGE ET GESTION DES DONNES.....	11
3.1. Le stockage des données multidimensionnelles.....	11
3.2. Le rôle du système de gestion des données.....	11
3.3. La structure du système de gestion (SGBD).....	12
3.3.1. Deux types de structures possibles.....	12
3.3.2. Exemples de SGBD relationnels.....	13
3.4. Les fonctions des SGBD et la manipulation des données.....	13
3.4.1. La manipulation des données descriptives...	13
3.4.2. Les manipulations spécifiques liées à l'attribut de localisation.....	14
3.5. Exemples de possibilités offertes par deux SGBD...	15
3.5.1. La base de données de l'archipel de Vanatu.	15
3.5.2. Le système TIGER.....	15

4. LE TRAITEMENT DES DONNEES DESCRIPTIVES EN RELATION AVEC LA VULNERABILITE DES NAPPES.....	16
4.1. Le traitement hydrogéologique et les processus physico-chimiques en cause.....	16
4.1.1. Les paramètres à représenter.....	16
4.1.2. L'élimination des polluants dans les eaux souterraines.....	17
4.1.3. Les processus physico-chimiques intervenant lors de la migration des polluants.....	19
4.2. Le traitement statistique des données.....	23
4.2.1. Les analyses statistiques descriptives.....	23
4.2.2. Les méthodes géostatistiques.....	24
4.3. Le traitement de l'image en vue de la restitution graphique.....	26
5. LES ORGANES DE SORTIE ET LES MODES DE RESTITUTION.....	29
6. LES ORIENTATIONS FUTURES DES SYSTEMES D'INFORMATION GEOGRAPHIQUE.....	30
6.1. Les avantages divers offerts par l'infographie....	30
6.2. Les croisements de cartes thématiques.....	30
6.3. Les systèmes experts.....	31
CONCLUSION.....	33
BIBLIOGRAPHIE.....	35

INTRODUCTION

DE LA NOTION DE CARTE A LA NOTION DE DONNEES LOCALISEES

Une carte représente un large ensemble d'informations. La localisation de chacune de ces informations dans l'espace est fixée graphiquement sur la carte. L'informatique offre des possibilités supplémentaires pour exploiter les informations contenues dans une carte, ou pour produire une nouvelle carte à partir d'informations de diverses origines. Dans les deux cas, le recours à l'informatique en cartographie, nécessite de considérer toutes les informations représentées sur une carte comme des données localisées.

Cet aspect spatial de l'information est essentiel, il induit des méthodes de traitement et de manipulation particulières aux données localisées. La localisation spatiale peut se doubler d'une localisation temporelle.

L'aspect spatial de l'information est associé à d'autres aspects concernant le thème de la carte, qui est ici la vulnérabilité des eaux souterraines aux polluants. Ces informations thématiques consistent, selon les cas :

- . en données brutes de terrain (exemples : profondeur de la surface piézométrique, épaisseur de l'aquifère),
- . en mesures de laboratoire (exemple : porosité d'échantillons de roche),
- . en données extraites d'une autre carte thématique (exemple : carte des transmissivités d'un aquifère) ou de fichiers de données,
- . en données calculées après un traitement plus ou moins complexe de données élémentaires (exemple : vitesse de migration d'un polluant dans une couche de terrain).

L'ensemble des opérations informatiques sur l'information localisée dans l'espace, allant de la saisie des données élémentaires en passant par le stockage, la gestion, le traitement et la restitution graphique de données localisées, s'intègre dans un système d'information géographique (M. Souris, 1986).

Cette note de synthèse présente l'ensemble des informations recueillies sur la cartographie de vulnérabilité des eaux souterraines à travers la bibliographie citée, en suivant le fonctionnement d'un tel système d'information géographique. Ce système, fictif, proposerait le maximum de possibilités actuellement offertes séparément, par les divers systèmes cités dans les ouvrages et articles consultés.

1. LE CHOIX DES DONNEES PARTICIPANT A L'ELABORATION DES CARTES DE VULNERABILITE.

Des choix doivent être précisés concernant les deux aspects de ces données,

- . d'une part, l'aspect spatial sur lequel s'exerce un certain nombre de contraintes, en relation avec la gestion et le traitement ultérieur de ces données, dites graphiques,
- . d'autre part, l'aspect descriptif de ces données, dont le choix est guidé par leur caractère de pertinence vis à vis du thème de la vulnérabilité.

1.1. LES CONTRAINTES CONCERNANT L'ASPECT SPATIAL DES DONNEES.

La localisation des points d'observation et des données doit être établie de façon précise par rapport à un système de référence géographique unique, quelle que soit l'échelle de travail.

En effet, selon l'échelle considérée, une ville pourra être considérée comme deux entités géographiques distinctes, un point sur une carte à grande échelle, ou une surface (c'est à dire un ensemble de points) sur une carte à plus petite échelle.

Ainsi, chaque entité géographique doit être définie par rapport à un système de référence unique, comme étant un ensemble de points. Le point étant l'élément de base de l'espace, chaque point doit être codé dans le cadre d'un modèle préétabli en vue de la création d'un index de points associés à des données géographiques (les coordonnées) et à des données descriptives.

L'espace est divisé en sous-ensembles suivant deux méthodes possibles :

- . le maillage qui est un découpage, à priori, indépendant des entités géographiques à représenter,
- . le recensement des sous-ensembles propres de l'espace, à savoir les points, les lignes, les zones constituant les entités géographiques.

Le système TC de l'ORSTOM (M. Souris, 1986) permettant de manipuler des données cadastrales, fonctionne suivant cette deuxième méthode : les zonages sont définis à posteriori, après avoir établi le schéma décrivant les ensembles d'objets à utiliser.

Cette division de l'espace en sous-ensembles permet d'en faire une description informatique. Le point est alors représenté par ses coordonnées, cartésiennes, cylindriques ou sphériques. La ligne est définie comme un ensemble de segments joignant des points appelés sommets. La zone est définie par ses contours, qui peuvent être représentés suivant deux méthodes :

- . le contour est un ensemble de courbes fermées, appelées boucles (chaque contour est alors stocké en mémoire deux fois : une fois pour chaque zone adjacente),

. ou bien, le contour est un ensemble de frontières ou lignes, c'est à dire d'arcs situés entre deux noeuds. Dans le système TC de l'ORSTOM précédemment cité, les points sont représentés en coordonnées sphériques relatives dans un système de références unique. Leur saisie se fait dans un module spécifique qui les numérote; ils sont pointés à la loupe à réticule. Un autre module effectue la saisie des lignes, elles sont numérotées, et des points sont numérisés en quantité suffisante pour suivre la courbure. Un autre module effectue la saisie des zones qui sont numérotées, avec suivi des frontières et indication du numéro de zone adjacente. Le recalage des données graphiques est réalisé grâce à trois points, qui ont été saisis sur chaque document à cet effet. Ces entités géographiques simples permettent de décrire graphiquement d'autres entités géographiques plus complexes, composées des éléments précédents . Il sera ensuite, possible de décrire les relations qui lient ces éléments, comme les flux, les migrations, les réseaux linéaires. On obtient alors, un graphe.

Toutes ces entités géographiques se décomposent en ensembles de points dont la description informatique est une application liant le numéro d'ordre du point dans l'index, aux coordonnées de ce point dans le plan ou dans l'espace.

1.2. LES DONNEES DESCRIPTIVES EN RELATION AVEC LE THEME DE VULNERABILITE DES NAPPES.

Il s'agit de données brutes, à l'exclusion de toute donnée de synthèse, même partielle.

Ces données peuvent être réparties en trois groupes :

- . les données concernant le site hydrogéologique étudié,
- . les données concernant les polluants,
- . les données concernant l'utilisation du site.

Cette classification est celle utilisée par W.A.M. Kerkhof et coll. (1987) dans la présentation de leur système d'évaluation automatisé de la pollution. Dans de nombreux travaux, la notion de vulnérabilité est moins précise; elle exprime les caractéristiques du site étudié par rapport à la pollution en général (A. Brewsmas et W. Van Duyvenbooden, 1987; R.S. Kerr, 1985).

D'autres travaux complètent cette approche en introduisant la notion de risque lié à des polluants donnés comme les pluies acides (M. Holmberg et coll., 1987).

Les problèmes liés à l'utilisation du site sont évoqués de façon qualitative par F.W. Schwartz et coll. (1982). Seuls, W.A.M. Kerkhof et coll. (1987) proposent une évaluation quantitative par classes.

1.2.1. Les données concernant le site.

L'étude la plus complète en est donnée par R.S. Kerr (1985) pour le système Drastic.

* Les données pédologiques :

Elles peuvent être extraites d'une carte des sols, et comprennent :

- . la nature du sol et, en particulier, le rapport calcaire / argile,
- . l'épaisseur du sol,
- . son taux de matières organiques,
- . sa granulométrie.

* Les données hydrogéologiques :

Elles peuvent être en grande partie, extraites de cartes hydrogéologiques, et complétées par des observations et des mesures de terrain. Ce sont :

- . les données de la zone non saturée :
 - + la nature des roches et des minéraux,
 - + l'épaisseur de cette zone,
 - + le degré de fracturation,
 - + le litage, susceptible de donner des couches servant de guides hydrauliques, ou d'autres types d'hétérogénéités,
 - + la conductivité hydraulique;
- . les données de la zone saturée :
 - + la nature des roches,
 - + leur porosité primaire,
 - + leur porosité secondaire, c'est à dire les diaclases et les fractures.

* Les données hydrodynamiques :

- . la profondeur et la pente de la surface piézométrique,
- . le gradient hydraulique,
- . la recharge de l'aquifère par unité de surface,
- . la mesure des forces de capillarité (à prendre en compte dans certains cas de pollution organique).

1.2.2. Les données concernant les polluants.

Le travail le plus complet est celui effectué sur les PCB par F.W. Schwartz et coll. (1982). Les paramètres chimiques et biochimiques considérés sont :

- . la nature du polluant et sa composition,
- . la concentration du polluant ou de ses différents éléments, dans l'eau,
- . ses caractéristiques physico-chimiques :
 - + sa densité,
 - + sa solubilité dans l'eau,
 - + sa viscosité,
 - + la tension d'interface polluant-eau, afin d'évaluer l'importance des phénomènes d'adsorption et le nombre de phases entre lesquelles se répartit le polluant en présence d'eau et de roches.
 - + le comportement chimique du polluant : son inertie, s'il est ionique ou non, s'il est hydrofuge ou non;
- . la présence éventuelle d'autres polluants sur le site, avec lesquels il y aurait possibilité d'interaction (les huiles et les détergents sont parmi les plus importants). Dans l'un de ces cas, il faut faire l'étude de la solubilité du

polluant dans l'huile et l'étude de la densité du nouveau mélange, ou l'étude de la variation des tensions d'interface polluant-eau en présence de détergents;

. les variations de pH ou de température; celles-ci pourraient entraîner des déplacements des équilibre chimiques entre phases.

1.2 Les données concernant l'utilisation du site et des alentours

Seuls W.A.M. Kerkhof et coll. (1987) les prennent en compte en tant que paramètres. Ils distinguent :

- . l'utilisation des terres sur le site et aux alentours,
- . l'utilisation des eaux de surface,
- . l'utilisation des eaux souterraines,
- . la nature et l'importance de la végétation sur le site, ainsi que son état.

Les données peuvent être extraites de plans d'aménagement, de cartes des sols et de la végétation. Elles peuvent être complétées et actualisées ou précisées par des enquêtes auprès des autorités locales ou sur le terrain.

Les données concernant l'utilisation des eaux souterraines ou des eaux de surface font appel à des paramètres biochimiques, tels que DBO (demande biochimique en oxygène), et biologiques (taux de bactéries coliformes permettant d'évaluer la pollution d'origine fécale). L'article de Coble (1981) traite de la distribution géographique de ces paramètres dans un cas précis de pollution par lagunage d'eaux usées.

2. LA SAISIE DES DONNEES

2.1. LA SAISIE DES DONNEES GRAPHIQUES.

Elle peut être manuelle au moyen d'une table à numériser, ou automatique suivant différentes méthodes (M. Souris, 1986).

2.1.1. La saisie sur table à numériser, ou digitaliseur.

Elle s'effectue en mode vecteur et s'applique aux cartes à implantation ponctuelle ou linéaire (réseau). Cette méthode offre la possibilité de passer de l'automatique au manuel. Ceci permet ainsi des contrôles interactifs qui autorisent une précision de 10 , pour saisir les coordonnées des points, des noeuds et des sommets.

Ce procédé est long, et par conséquent coûteux; il faut compter 3 à 5 jours à un opérateur pour saisir une carte de 500 à 800 zones. Il a été utilisé pour la carte des sols du Nord-Cameron (P. Brabant, 1986); dans ce cas, il a permis d'effectuer la planimétrie des zones avec une précision de 1% .

2.1.2. La saisie par balayage.

Elle s'effectue en mode raster au moyen d'un scanner ou d'un densitomètre à balayage.

Elle permet la saisie automatique de documents suivant une grille de niveau par balayage horizontal, ligne après ligne. Aussi, nécessite-t-elle un algorithme de passage d'une représentation maillée correspondant au mode raster à une représentation linéaire correspondant au mode vecteur.

Le document à saisir ne doit contenir que l'information à saisir, à l'exclusion des noms, des symboles ou des fonds de carte. Tout autre signe que les éléments à saisir engendre des parasites.

Cette méthode n'est pas adaptée à la saisie de points isolés, mais à celle des zones.

Elle a l'avantage d'être rapide, mais elle nécessite un traitement ultérieur manuel pour les vérifications topologiques.

2.1.3. La saisie par suivi automatique du trait.

Elle s'effectue au moyen d'une diode ou d'une caméra laser, et s'applique aux cartes à implantation linéaire. Elle correspond à un compromis entre les deux précédentes. Elle nécessite également un document ne comportant que l'information à saisir, et des vérifications topologiques ultérieures.

2.1.4. Les corrections et modifications accompagnant la saisie.

Ces opérations nécessitent un éditeur graphique pour afficher la zone à modifier ou pour retrouver un arc, ainsi que des modules de modification interactifs, qui conservent l'intégralité des relations topologiques.

Ces modules de modification interactifs doivent permettre de vérifier la fermeture des zones et d'éliminer les problèmes graphiques, tels que connexité, simplicité et intersection d'arcs. Ceci est effectué par élimination des points superflus, et par lissage des courbes, afin de réaliser le filtrage des arcs.

Ils doivent également assurer les raccordements de trait entre les différentes cartes saisies lors de leur assemblage.

Le système TIGER de l'unité infographique de l'ORSTOM, cité en exemple, possède un module de saisie graphique par digitaliseur assurant ces fonctions. Ce module gère les problèmes de juxtaposition des fenêtres de saisie, qui sont définies par l'opérateur, ainsi que leur reprise en cours de manipulation.

La numérisation se fait à partir d'une minute de terrain à trait perdu sur support stable, à trait régulier et contours fermés.

Les procédés de lissage des courbes, utilisés pour la réalisation de l'atlas structural de la Suisse (K. Brassel et coll., 1985) sont expliqués en détail dans l'EURO CARTO IV Seminar.

2.2. LA SAISIE DES DONNEES DESCRIPTIVES.

La saisie des données graphiques a pour résultat une liste d'unités cartographiques, chacune dotée d'un numéro et d'une clé de descripteurs. A chacune de ces unités cartographiques sont associées des données descriptives, dont la saisie s'effectue après avoir sélectionné l'unité cartographique correspondante. La sélection a lieu, soit en indiquant le numéro de cette unité, soit en pointant cette unité sur un écran de visualisation au moyen d'un stylet ou d'une souris.

3. LE STOCKAGE ET LA GESTION DES DONNEES

3.1. LE STOCKAGE DES DONNEES MULTIDIMENSIONNELLES.

Il est effectué sur disque ou sur bande. Les données sont organisées en fichiers. Leur volume dépend de la technique de structuration de l'information choisie : le mode vecteur ou le mode maillé.

* Les données structurées en mode vecteur occupent un volume réduit, mais en contrepartie, leur saisie et leur recherche sont plus longues. Ainsi, pour stocker les données d'une carte de quelques centaines de zones, saisies avec une précision de 0,5mm, il faut une place en mémoire de 100 à 500 ko (M. Souris, 1986).

* Les données structurées en mode maillé ou raster occupent un volume 10 fois plus important, ce qui oblige à recourir à des techniques de compactage.

Il existe plusieurs méthodes de compactage des données :

- . le codage en chaîne ou incrémental, tel que :
 - + le code de Freeman (il ramène l'occupation mémoire à 3 bits au lieu de 20 pour chaque point stocké),
 - + le codage à 4 directions par incrément de courbure,
 - + le codage à 4 directions par vecteurs de type Freeman, mais à norme variable; c'est le plus efficace pour la compacité;
- le codage en chaîne engendre des difficultés dans les manipulations graphiques d'un élément codé;
 - . le codage en squelette,
 - . la représentation par quadrees; elle correspond à une structure arborescente avec découpage de l'espace en 4 aires. Ce découpage peut être arbitraire en 4 zones égales, ou récursif et dépendant du nombre d'objets à représenter dans chaque découpage. C'est la méthode la plus intéressante, à la fois pour sa compacité et pour le fait qu'elle est bien adaptée aux opérations graphiques et au calcul des propriétés géométriques.

3.2. LE ROLE DU SYSTEME DE GESTION DES DONNEES.

Les fichiers de données graphiques et descriptives sont gérés par un système de gestion de base de données (SGBD) sans intervention manuelle, qui réalise les interfaces entre les mémoires de stockage et les usagers.

Le but du SGBD est d'assurer la sécurité, la cohérence et la non redondance des données stockées. Pour cela, il doit y avoir indépendance logique entre la structure de données d'une application et la structure de données modélisant le monde réel.

Ainsi, on distingue 3 niveaux de description des données :

- . le niveau externe, où les données sont telles qu'elles sont perçues par l'utilisateur;
- . le niveau conceptuel, qui est un modèle entités-associations; il comprend :
 - + des données élémentaires définissant les attributs ou caractéristiques des objets,
 - + des données composées correspondant aux objets réels,
 - + des données composées décrivant des associations d'objets réels; elles contiennent alors des attributs provenant d'autres entités; ce modèle doit respecter des règles de cohérence entre les attributs d'une entité ou d'une association;
- . le niveau interne correspond à la structure de stockage, c'est à dire aux fichiers et aux chemins d'accès aux données.

Le SGBD possède donc 3 couches successives de fonctions, correspondant à ces 3 niveaux :

- . le système de gestion de fichiers proprement dit, assure la liaison entre le disque (organisation physique) et les fichiers (organisation logique);
- . le SGBD interne gère l'accès aux données (placement, assemblage et liens entre les données et les schémas);
- . le SGBD externe gère la mise en forme des données pour les usagers en assurant l'interprétation et l'optimisation des requêtes.

3.3. LA STRUCTURE DU SGBD.

3.3.1. Il existe deux types de structures possibles pour le SGBD :

* La structure hiérarchique correspond à des codes emboîtés; il s'agit d'une structure figée.

* La structure relationnelle, beaucoup plus souple, est actuellement la plus utilisée. Entre les données graphiques et descriptives sont définies des relations associant des attributs, qui décrivent les objets des entités ou associations.

Ces relations doivent éviter les redondances. Elles établissent des dépendances entre les attributs :

- . il y a dépendance fonctionnelle entre A et B, si à toute valeur de A correspond une valeur unique de B;
- . il y a dépendance multivaluée, si deux ensembles d'attributs indépendants sont, chacun, corrélé à un troisième;
- . les dépendances de jointure décrivent les relations entre sous-ensembles d'attributs d'une relation.

3.3.2. Exemples de SGBD relationnels.

* Le SGBD appliqué à l'archipel de Vanatu (xxxxxxxx ORSTOM , 1986) gère un ensemble de données pédologiques (tirées de cartes au 1 / 50 000), géomorphologiques, géologiques et de végétation (tirées de cartes au 1 / 100 000) et administratives (tirées de cartes au 1 / 100 000 , telles que les aires d'état civil, les aires de recensement, les groupes de peuplement).

Cinq relations ont été définies entre les données :

- . une relation zonale définissant le découpage cartographique;
- . la relation 1 concerne les sols, son seul attribut est la pédologie;
- . la relation 2 concerne la morphologie, elle a deux attributs : la géomorphologie et la géologie;
- . la relation 3 concerne la végétation, elle n'a qu'un attribut;
- . la relation 4 concerne l'administration, elle a 6 attributs , parmi lesquels les 4 derniers sont quantitatifs :
 - + les aires d'état civil,
 - + les aires de recensement,
 - + le nombre de ménages,
 - + le nombre de résidents,
 - + les groupes de peuplements,
 - + le nombre de lieux-dit.

* Le SGBD de manipulation de données cadastrales TC (M.Souris, 1986) compte quatre relations concernant :

- . le parcellaire, c'est à dire l'utilisation des sols; elle est tirée du cadastre;
- . le bâti; cette relation relie les constructions du cadastre aux descripteurs présents sur le terrain;
- . les différents réseaux, aériens (électricité) et souterrains (eau, gaz, électricité et assainissement);
- . les permis de construire.

3.4. LES FONCTIONS DES SGBD ET LA MANIPULATION DES DONNEES.

Aux fonctions classiques d'interrogation, mise à jour, insertion et calculs des SGBD s'ajoutent des manipulations spécifiques à l'attribut de localisation.

3.4.1. La manipulation des données descriptives.

Elles sont toutes basées sur des opérations d'algèbre relationnelle et comprennent :

- . l'extraction, la sélection et la mise en relation d'attributs descriptifs; ces manipulations permettent d'interroger la base et font appel aux opérations de :
 - + projection d'une relation sur certains de ses attributs,
 - + restriction d'une relation par une qualification,
 - + jointure de deux relations selon une qualification;

- . l'intégration de données dans la base; entre autres, d'images numériques de type satellitaire; cela nécessite un recalage avant la mise en relation de cette information avec l'information correspondante de la base;
- . la structuration, l'archivage, la recherche, la mise à jour et les modifications de données descriptives;
- . la création d'attributs de nouvelles variables par:
 - + des calculs numériques, statistiques ou logiques (union, différence, produit cartésien de deux relations),
 - + des classifications,
 - + des techniques d'analyse des données;
- . le tri et l'impression des résultats.

3.4.2. Manipulations spécifiques liées à l'attribut de localisation.

* L'opération de restriction spatiale réalise une sélection d'objets par rapport à la localisation. Cette opération permet les manipulations suivantes :

- . l'extraction et la sélection de feuilles géographiques,
- . la définition de fenêtres en coordonnées géographiques, ou par rapport à d'autres objets géographiques (zone tampon, couloir, zone d'influence).

* La mise en relation géographique comprend :

- . les croisements de zones;
- . les croisements de points et de zones;

ces deux types de manipulations sont obtenues grâce à l'opération de jointure spatiale, qui est une mise en relation de deux objets par rapport à leur localisation respective;

- . la création de nouveaux objets graphiques ayant des valeurs descriptives provenant de plusieurs relations différentes;
- . le dénombrement d'objets ponctuels situés dans des zones pré-sélectionnées.

* L'agrégation graphique de zones résulte d'une classification sur les attributs descriptifs. Cela nécessite la suppression des frontières de zones adjacentes appartenant à une même classe.

* La création d'éléments par interpolation, tels que les centroïdes et les zonages s'effectuent à partir de points valués, de courbes d'isovaleurs, et de modèles numériques de terrain.

* Les mesures géométriques suivantes sont possibles :

- . la distance entre 2 points;
- . la distance d'un point à une ligne ou à une zone;
- . la distance le long d'une ligne;
- . le calcul des superficies, des périmètres.

* Parmi les calculs d'optimisation, l'un des plus utilisé est celui du plus court chemin.

* Les calculs en trois dimensions, tels que la visibilité, la réflexion ou les écoulements dépendent de la projection utilisée. Aussi, les systèmes d'information géographique présentent-ils généralement un choix de projections de travail et de restitution.

3.5. EXEMPLES DE POSSIBILITES OFFERTES PAR DEUX SGBD PRECEDEMMENT CITES.

3.5.1. La base de données de l'archipel de Vanatu.

Son module d'interrogation permet la définition d'une fenêtre dans l'espace géographique de la base, à savoir 166 à 172 de longitude pour 12 à 22 de latitude sud. La fenêtre est déterminée par les coordonnées du point inférieur gauche et celles du point supérieur droit. La sélection s'opère sur les éléments pris en compte dans la base, leurs valeurs étant agrégées en classes. Des informations supplémentaires dans la fenêtre choisie peuvent être obtenues sur une seule relation sous forme d'une image complémentaire.

3.5.2. Le système TIGER

La sélection de la fenêtre d'étude peut se faire suivant quatre modalités :

- . directement, en coordonnées géographiques par son point supérieur droit (colatitude minimale, longitude maximale) et son point inférieur gauche (colatitude maximale, longitude minimale);
- . ou en indiquant la feuille; la fenêtre correspond alors, au domaine géographique de la feuille;
- . ou en indiquant le point inférieur gauche et la résolution du pixel en mètre;
- . ou en indiquant un certain nombre de valeurs d'attributs d'une relation; la fenêtre sera alors le domaine minimal comprenant l'ensemble des objets répondant aux valeurs demandées.

L'intégration des images numériques de type satellitaire est facilitée grâce à la transformation de la structure vectorielle des données graphiques en structure matricielle, sous forme d'images numériques de pixels associées à un fichier descriptif. Pour intégrer les données satellitaires , il suffit de faire coïncider les paramètres des deux matrices de pixels.

L'intégration de ces données images permet la création ou la modification de relations. L'extraction d'information à partir de ces données images serait facilitée par les techniques de l'intelligence artificielle. Cette phase est actuellement en projet. La base de faits d'un tel système expert serait constituée par les résultats d'opérateurs de reconnaissance de forme sur l'image, et par les résultats de l'interrogation des données existantes dans la base.

4. LE TRAITEMENT DES DONNEES DESCRIPTIVES EN RELATION AVEC LA VULNERABILITE DES NAPPES.

Le traitement des données descriptives, en relation avec la vulnérabilité des nappes, comprend plusieurs phases.

Le traitement hydrologique a pour objectif, l'estimation des processus physico-chimiques, qui entrent en jeu, lorsque des polluants menacent des nappes souterraines, s'y déplacent ou y évoluent.

Le traitement statistique des données permet de décrire, d'analyser les données hydrologiques, ou de restituer la distribution spatiale des paramètres caractéristiques de la vulnérabilité des nappes, en vue d'établir une carte thématique.

Le traitement de l'image vient en dernier lieu. Son objectif est de fournir la restitution graphique qui correspondra à la meilleure figuration de la réponse attendue par les usagers.

4.1. LE TRAITEMENT HYDROGEOLOGIQUE ET LES PROCESSUS PHYSICO-CHIMIQUES EN CAUSE DANS LA VULNERABILITE DES NAPPES AUX POLLUANTS.

4.1.1. Les paramètres à représenter.

La représentation sur une carte thématique de la vulnérabilité des nappes aux polluants repose sur les deux concepts suivants (J.L. Anderson, 1987).

* La capacité de nettoyage des sols est un paramètre destiné à exprimer la capacité des sols à éliminer un polluant donné. L'unité préconisée est la masse de polluant éliminée par unité de volume de sol.

Le rôle de la matière organique apparaît comme primordial, dans les processus de filtration, de complexation, d'adsorption, de précipitation, de dégénérescence biologique, qui entraînent l'élimination partielle du polluant des eaux souterraines. Aussi, plusieurs auteurs s'intéressent-ils préférentiellement à la couche de 0 à 0,5m de profondeur, car elle est supposée contenir la plus grande partie de la matière organique.

J.L. Anderson ne donne pas de méthode pour estimer ou calculer ce paramètre. Par contre, ce concept a été appliqué par M. Holmberg et coll. (1987) à la pollution par les pluies acides.

La capacité de nettoyage du sol devient alors, la capacité de neutralisation du sol; elle est exprimée en fonction de sa teneur en cations de base de la couche supérieure de sol, de 0 à 0,5m de profondeur. Interviennent également dans cette évaluation :

. l'épaisseur du sol (appréciée suivant deux classes : comprise entre 0 et 0,5m, ou supérieure à 0,5m de profondeur)

. la granulométrie du sol (appréciée suivant cinq classes granulométriques).

C'est également la profondeur de 0,5m, qui a été retenue par B.L. Velikov et Y. Al Mohammed (1987), pour prélever les échantillons de sol destinés à des essais de laboratoire sur la migration des pesticides en milieu poreux.

* Le second concept est la capacité de recharge de l'aquifère. Elle est exprimée par le temps de séjour moyen des eaux dans l'aquifère, qui est égal au volume d'eau de l'aquifère, divisé par le taux de recharge.

Au niveau de l'aquifère, c'est essentiellement le renouvellement des eaux, qui va permettre d'éliminer le polluant en l'entraînant.

L'application de ce second concept, aux pluies acides, conduit à estimer la capacité de neutralisation de l'aquifère, qui dépend non seulement du potentiel de recharge annuel de l'aquifère, mais aussi de la solubilité ou altérabilité des minéraux de l'aquifère.

4.1.2. L'élimination des polluants dans les eaux souterraines.

D'autres auteurs ont tenté d'évaluer l'élimination des substances polluantes dans les eaux souterraines, soit par des essais de laboratoire , soit par des études de terrain .

* Les essais de laboratoire (B.L. Velikov et Y. Al Mohammed, 1987) ont été effectués sur des échantillons de sable pris à 0,5m de profondeur, et à 4m de profondeur. La taille des échantillons n'est pas précisée, mais ces échantillons sont supposés représenter un milieu poreux, hétérogène. Les polluants étudiés sont réputés stables, il s'agit des pesticides suivants :

. l'atrazine (2-chlor-4éthylamine-6isopropylamine-sym-triazine);

. le saturn (S-/4-chlorobenzyl/-N-N-diethylthiocarbamate);

. la phenyltrothione (0,0-dimethyl-0-/3-methyl-4nitrophenyl/thiophosphate).

L'élimination des polluants au cours du temps est approchée par la formule :

$$dC / dt = g.C^n$$

où C est la concentration d'un polluant à l'instant t, et g son taux d'élimination.

Grâce à un programme informatique, on recherche l'ordre n, de 0 à 6, qui fournit la meilleure adéquation (la valeur du coefficient de corrélation R, est alors maximale) avec les résultats de l'essai, dont la durée est de 48 heures. Puis, on cherche g, la constante du taux d'élimination, et T(1/2), la durée au bout de laquelle la concentration du polluant a diminué de moitié. Un autre programme détermine un coefficient de dispersion :

$$Kd = (V/m).(Co/Ci - 1)$$

dépendant du temps, car l'équilibre thermodynamique n'est pas atteint au bout de 48 heures. L'interprétation des données des

essais est alors basée sur un modèle de migration linéaire à deux variantes : l'une d'élimination par dispersion, l'autre d'élimination par dispersion-diffusion.

* L'étude de terrain (C.W. Welby, 1987) a porté sur l'élimination de polluants sur le site de Raleigh (Caroline du Sud, U.S.A.), pendant une durée d'observation de 4 ans. Les polluants en cause, étaient des herbicides, des pesticides et des solvants ordinaires, enfouis sur le site entre 1969 et 1980; cette période est supposée assez longue pour pouvoir considérer le système comme étant en équilibre vis à vis de la source d'élimination.

Les roches de l'aquifère sont des sables fins et des silts argileux, dont le diamètre moyen est 3,5. Ces sédiments résultent de l'altération sur place de la roche-mère (gneiss et schistes), dont ils ont conservé la foliation. Le plongement de la foliation n'est pas constant sur le site, ce qui ajoute encore à l'hétérogénéité du milieu.

La porosité moyenne est environ de 43%.

Le taux de matière organique est inférieur à 0,01%.

La conductivité hydraulique moyenne est de l'ordre de 0,0001cm/s; le gradient hydraulique est orienté vers le NW, mais localement des niveaux plus perméables, parallèles à la schistosité, peuvent contrôler les écoulements et le déplacement des polluants.

Les déchets chimiques ont été enterrés sur le site dans des tranchées de 3m de profondeur, et de 1 à 2m de large. Celles-ci constituent donc, des sources ponctuelles multiples.

Le tétrachlorure de carbone a été choisi parmi les 15 composés étudiés, pour illustrer la méthode de travail. Les données des analyses de ce polluant, réalisées à partir de prélèvements d'eau dans les puits, pendant quatre ans, ont été traitées par leur moyenne géométrique. Elles vérifient la relation :

$$C = C_0 \cdot e^{-uS}$$

avec C : concentration à la distance S de la source de pollution

C₀ : concentration du polluant à la source

u : "constante d'élimination" relative à l'élimination du polluant le long du profil choisi.

Cette relation permet de déterminer, perpendiculairement aux équipotentielles, la vitesse de migration de la pollution. Elle est calculée à partir de la distance parcourue par le front de pollution, depuis la source de pollution, dans un intervalle de temps donné (pour cette étude, elle est de 30m en 15 ans). Le front de pollution est caractérisé par le fait que la concentration est nulle au-delà; c'est à dire que C tend vers 0 sur cette limite.

Cette équation permet également de calculer la distance qu'aura parcourue le front de pollution à un instant ultérieur, et la concentration de polluant attendue à une distance fixée de la source de pollution.

La forme exponentielle de la relation montre les risques qu'il peut y avoir à tenter de déterminer le coefficient d'élimination à une très faible distance de la source de pollution, ou au bout d'un laps de temps trop court.

L'importance de cette étude réside dans le fait qu'il s'agit d'une étude de terrain effectuée dans des conditions réelles, en milieu hétérogène aussi bien en ce qui concerne la nature lithologique des couches que leur structure, et qui a porté sur de nombreux agents polluants (15). L'étude a duré 4 ans, couvrant ainsi les variations hydrodynamiques de la nappe durant 4 cycles hydrologiques. Les effets de la pollution suivie, se sont étalés et stabilisés durant une période assez longue : plus de 15 ans. Etant donné que les déchets sont enfouis à 3m de profondeur, l'étude fait abstraction du rôle autoépurateur du sol.

Pour ces raisons, les résultats de cette étude sont fiables; malheureusement, son coût en temps, en équipement (nombreux piézomètres de surveillance), et en analyses chimiques (15 composés organiques polluants analysés en chromatographie, à raison d'un échantillon ou 2 par semaine pendant 4 ans sur chaque piézomètre) rend inaccessible cette méthode dans bien des cas de pollution.

D'autres approches tentent d'évaluer les phénomènes responsables de l'élimination de façon moins globale, non plus à l'échelle d'un site, mais à l'échelle des couches multimétriques et des grains des sédiments, ce qui nécessitent des investissements moins lourds.

4.1.3. Les processus physico-chimiques intervenant lors de la migration des polluants dans les eaux souterraines.

4.1.3.1. Le comportement des polluants dans l'eau (solubilité, miscibilité et nombre de phases).

Souvent, les problèmes concernant la répartition du polluant en différentes phases sont posés en terme de solubilité dans l'eau. Seul le devenir de la partie soluble est pris en compte, car son comportement hydrodynamique est plus facile à caractériser et à modéliser que celui de la partie non miscible. Le comportement de la partie non miscible est éludé en considérant qu'elle est retenue par adsorption (K.U. Weyer, 1987).

Pourtant, c'est souvent la phase non miscible, de densité supérieure à celle de l'eau, qui est responsable de la forme et de l'extension du domaine pollué.

L'importance des phases non solubles dans l'eau est démontrée par l'étude d'un déversement de PCB sur le site de Regina au Canada (F.W. Schwartz et coll., 1982). Le comportement des 3 phases : aqueuse dissoute, adsorbée et huileuse non miscible, a été envisagé. Le produit en cause est l'Inerteen (mélange de 70% de PCB : Arachor, et 30% de TCB : tri et tétrachlorobenzène); il se présente comme un liquide inerte, dense, visqueux, peu soluble (concentration maximale dans l'eau : 0,04mg/kg), fortement adsorbé sur la matière organique, mais faiblement adsorbé sur les minéraux.

Le rôle joué par la structure géologique des couches concernées, intervient à ce niveau de façon décisive.

L'existence de fractures interdit d'assimiler la couverture de dépôts glaciaires argileux à un milieu imperméable, ou peu perméable; car, les observations ont montré que ces fractures constituent un important facteur d'expansion de la pollution. Aussi, les coefficients de partage du polluant entre les différentes phases, exprimés en terme de masse de milieu poreux homogène, ne sont plus directement applicables. Pour évaluer la probabilité d'avoir 2 phases liquides ou la possibilité de migration de la phase dissoute dans un milieu poreux non fracturé, il faut se ramener au niveau des pores et des interfaces entre solides et liquides, et s'intéresser aux tensions d'interface, à la mouillabilité et aux forces capillaires.

La tension d'interface de l'Inerteen avec l'eau, a été estimée comme intermédiaire entre celle du benzène (35,0) et celle du carbone (47,6). Cet ordre de grandeur indique que l'Inerteen et l'eau vont constituer 2 phases dans un milieu poreux, et que l'Inerteen aura fort peu tendance à adhérer aux grains des minéraux ou à se mélanger à l'eau.

Partant de là, les cas suivants ont été envisagés :

- . sous la surface piézométrique, l'Inerteen est une phase non mouillante par rapport à l'eau; et le PCB occupe l'espace restant à l'intérieur des pores, à condition d'avoir pu vaincre les forces de capillarité pour pénétrer à l'intérieur de ces pores;

- . au-dessus de la surface piézométrique dans la zone où les teneurs en eau restent fortes, l'eau reste la seule phase mouillante; l'Inerteen et le gaz occupent donc l'espace restant à l'intérieur des pores;

- . au-dessus de la surface piézométrique, près de la surface du sol, où les teneurs en eau sont assez basses, l'Inerteen peut mouiller le milieu poreux et spontanément remplacer le gaz.

Dans le cas réel d'un déversement de PCB, les systèmes sont beaucoup plus complexes; en effet, la présence d'autres produits est facteur de complication. Les TCB, par exemple, étant plus solubles dans l'eau que les PCB, vont faire augmenter les concentrations des PCB dans l'eau jusqu'à une valeur inconnue. Les huiles minérales qui ont également été déversées sur le site, influencent le schéma de migration; en effet, les PCB sont fortement solubles dans les huiles, avec une densité inférieure à 1; par conséquent, leurs direction et quantité de déplacement vont différer considérablement de ceux de l'Inerteen pure.

Ainsi, les observations sur le site ont fait apparaître les points suivants.

* Grâce aux fractures, l'essentiel des PCB en phase huileuse traverse la couche d'argile (où les vitesses à travers les pores ne sont pourtant que de quelques mm /an) et gagne la couche de silt sous-jacente. Une petite partie seulement des PCB, la fraction soluble dans l'eau, migre à travers le silt vers la surface piézométrique de l'aquifère principal holocène. Seule cette fraction minime est soumise à l'atténuation des concentrations par diffusion matricielle.

* Dans l'aquifère perché, le flux latéral a favorisé l'extension latérale des PCB et TCB jusqu'à 90m du point de déversement. Les très fortes concentrations de PCB observées dans le sol, indiquent que le transfert s'est fait sous la forme d'une phase pure d'Inerteen.

Cet exemple montre la nécessité de prendre en considération toutes les phases du polluant et d'étudier leur comportement en relation avec les structures géologiques.

4.1.3.2. Le temps de transit du polluant à travers la zone non saturée.

La zone non saturée est considérée comme un empilement de couches homogènes dans le modèle multicouche (Y. Bachmat et M. Collin, 1987).

La durée de traversée du polluant à travers une séquence de N couches lithologiques homogènes est :

$$T = \text{somme de } 1 \text{ à } N \text{ de } (L/V)_i$$

avec L : épaisseur de la couche homogène i,

et V : vitesse moyenne verticale du polluant à

travers la couche i.

La vitesse moyenne du polluant peut être évaluée suivant 3 types de modèles d'écoulement :

. le modèle "piston-flow" suppose que le polluant se déplace à la même vitesse que l'eau; dans ce cas :

$$V_i = q / h_i$$

avec q : débit spécifique vertical de l'eau d'infiltration

et h_i : taux d'humidité de la couche;

. le modèle "advection-dispersion" de Bear (1979), qui suppose que le polluant est transporté à la vitesse moyenne de l'eau, et dispersé selon les variations des vitesses des particules individuelles d'eau; le temps de transit devient :

$$T = \text{somme de } 1 \text{ à } N \text{ de } (h_i \cdot L_i / q - 16\alpha l \cdot (\text{racine de } (1 + L / 8\alpha l) - 1))$$

avec αl : dispersivité longitudinale du polluant à un taux d'humidité donné h_i ;

. le modèle "vitesse spécifique du polluant" suppose que le polluant se déplace à sa propre vitesse, indépendante de celle du vecteur eau; ainsi, dans les cas d'exclusion ionique, la vitesse de l'anion peut être supérieure à celle de l'eau (H. Gvirtzman et coll., 1986).

4.1.3.3. La fonction de dispersion de la phase soluble en milieu poreux saturé.

Elle est établie pour des solutions salines solubles en milieux aqueux cheminant dans un sédiment poreux, à l'origine homogène, isotrope, saturé. Les sédiments n'étant qu'exceptionnellement homogènes et isotropes, cette fonction de dispersion a été aménagée suivant différents modèles, pour tenter de la rendre applicable à des milieux réels complexes, hétérogènes ou anisotropes (J-B. Poulet, 1978).

* Le modèle multicouche de Marle assimile les sédiments à un milieu stratifié dont les différentes couches seraient homogènes.

* Le modèle de Roschal (1978) est un modèle déterministe des écoulements, assimilant les sédiments à des blocs hétérogènes ou à des couches semi-perméables répartis dans une matrice homogène perméable; la migration des polluants s'effectuerait par convection dans les zones perméables et par diffusion dans les zones semi-perméables.

Ces modèles analytiques seraient difficilement applicables à un milieu réel, selon J.B. Poulet, qui montre que la répartition des particules argileuses à l'intérieur des différentes couches, a une importance supérieure à celle de la vitesse de pores. En effet, les mécanismes de sorption réversibles ou irréversibles par les particules argileuses sont du même ordre de grandeur que les phénomènes dus aux mécanismes de dispersion.

* Le modèle de Gehlar (1978) est un modèle stochastique de macrodispersion, applicable à un milieu stratifié. Ce type de modèle a inspiré différentes études (G. Mathéron et coll., 1980 ; L.W. Gelhar et coll., 1979).

* Le modèle proposé par A. Dieulin (1980) s'inscrit dans la ligne des modèles stochastiques. Il s'applique à un milieu sédimentaire alluvial anisotrope, où la perméabilité présente des variations non négligeables au sein des couches assurant le cheminement préférentiel des écoulements.

La fonction de dispersion est :

$$K(t) = uP \cdot g(x)$$

avec u : vitesse moyenne de pores du milieu,

p : proche de la valeur $3/2$,

g : fonction intrinsèque de dispersion,

x : distance parcourue par le flux polluant.

La fonction $g(x)$ est calculée à partir des résultats des essais de traçage effectués sur le terrain. Ces traçages doivent couvrir toute la gamme des régimes d'écoulement probables (écoulement radial convergent ou radial divergent, ou écoulement plan).

4.1.3.4. Les interactions entre le polluant et son environnement en zone non saturée.

Ces interactions peuvent être de différente nature : adsorption, prélèvement par les racines des plantes, complexation, précipitation, filtration, et éventuellement dégénérescence due aux processus biologiques, chimiques, ou physiques.

Ces interactions entraînent un retard dans le transit du polluant à travers la zone non saturée. Ce retard est pris en compte en remplaçant le débit spécifique vertical de l'eau d'infiltration q , par le rapport q / R_d (Jackson et al., 1980).

La valeur R_d est le facteur retard, calculé de la façon suivante :

$$R_d = 1 + (db \cdot K_d / h_i)$$

avec db : densité de l'absorbant,
et K_d : coefficient de distribution entre la partie adsorbée et la partie dissoute du polluant.

4.2. LE TRAITEMENT STATISTIQUE DES DONNEES.

Le traitement statistique des données peut être appliqué, soient aux données brutes de terrain, soient aux paramètres calculés à la suite du traitement hydrogéologique des données.

Les traitements statistiques de données visent deux types d'objectifs :

- . soit décrire et analyser la distribution des données, afin d'évaluer les processus en cause; ces méthodes d'analyse vont des plus simples, analyses univariées, aux plus complexes, analyses multidimensionnelles;
- . soit restituer les valeurs des données en tout point d'un domaine de l'espace (ou du temps), à partir d'un réseau de points d'échantillonnage; ce sont les méthodes géostatistiques.

4.2.1. Les analyses statistiques descriptives.

Nombreux sont les logiciels de gestion de base de données à proposer des modules de calculs statistiques.

Le SGBD du système TIGER déjà cité, offre un tel module, nommé STAT. Ce module est appellable à n'importe quel moment de l'interrogation, ce qui permet de modifier la suite de l'interrogation en fonction des résultats présentés par les calculs statistiques. Par exemple, le choix des seuils de classification peut être effectué après avoir étudié la description statistique des paramètres de la classification. Les traitements statistiques sont univariés (calcul des moments, des histogrammes) ou bivariés (corrélations et régressions).

Le SGBD SAS (Ph. Waniez, 1986) non seulement :

- . effectue des traitements univariés (procédure UNIVARIATE),
 - . calcule les minimum, maximum, moyenne et écart-type (procédure MEANS),
 - mais aussi :
 - . travaille sur les variables centrées réduites (procédure STANDARD),
 - . traite les corrélations avec les coefficients de Pearson et Spearman (procédure CORR),
 - . calcule les régressions et donne le tableau des résidus stockés (procédure REG),
 - . effectue des analyses de surfaces de tendance avec formulation de modèles polynomiaux (jusqu'à l'ordre 4).
- Ce SGBD possède également plusieurs procédures d'analyse multidimensionnelle à n variables :
- . l'analyse en composantes principales (PRINCOMP),

- . l'analyse factorielle (FACTOR),
- . la classification automatique de type ascendante hiérarchique (CLUSTER),
- . l'analyse déterminante,
- . les nuées dynamiques.

Ces traitements peuvent être enchaînés au gré des besoins de l'utilisateur et des résultats apparaissant en cours de traitement.

Exemple d'enchaînement :

1. transformation en rang des variables,
2. analyse en composantes principales (ACP),
3. classification hiérarchique des observations sur le plan des facteurs I et II de l'ACP,
4. partition de l'ordre de classification obtenu précédemment en 5 classes.

4.2.2. Les méthodes géostatistiques.

4.2.2.1. Les conditions d'application de ces méthodes.

Ces méthodes reposent sur la théorie des variables régionalisées, applicables à des études de pollution dont la source est étendue au niveau régional (M.C. Rang et coll., 1987), comme la pollution des eaux souterraines par les dépôts atmosphériques, ou par l'inondation des sols par des eaux de surface polluées. Parmi ces types de pollution figurent celles engendrées par les épandages d'engrais, de pesticides, les eaux de ruissellement des chaussées et autoroutes, la contamination par un fleuve pollué, des nappes situées sur son tracé. Par contre, les cas de pollutions ponctuelles telles que les décharges domestiques ou industrielles, les décharges de voitures, les inondations de sites industriels de dimensions restreintes ne peuvent faire l'objet de ce type de traitement. En effet, pour que les méthodes géostatistiques soient applicables, la condition mathématique suivante doit être respectée; il faut que les variations de concentration des polluants soient continues et obéissent à une distribution statistique normale.

L'étude présentée concerne la pollution des sols à Eijsden (RFA). La pollution résulte d'apports diffus de métaux lourds (Ba, Zn, Pb, Cd) par 3 sources spécifiques :

- . les dépôts atmosphériques dûs aux cheminées d'usine de colorants;
- . par les eaux de lessivage des chaussées qui contiennent des métaux lourds dans leurs cendres de soubassement;
- . par les eaux d'inondation et les boues de la Meuse, drainant les déchets miniers et ceux des industries métallurgiques belges.

4.2.2.2. Les contraintes sur l'échantillonnage.

L'échantillonnage des sols doit obéir à certaines règles pour que soient respectées les conditions mathématiques initiales de continuité et de normalité de la distribution des données.

Les points de prélèvement sont situés à plus de 3km de l'usine de colorants, et à plus de 50m des routes dont les soubassements sont sources de pollution.

Afin que soit respectée l'homogénéité des données, les points d'échantillonnage ont tous été choisis dans des lopins cultivés et des jardins maraîchers, dont les modes de cultures diffèrent peu. Les terrains cultivés ont été choisis comme lieu de prélèvement pour représenter la pollution des sols, à cause de l'importance que revêt leur qualité pour la santé des populations.

En chacun des 62 points d'échantillonnage, 15 prélèvements ont été effectués, répartis sur une surface de 50m², à une profondeur de 20cm dans le sol.

4.2.2.3. La base des calculs géostatistiques.

La structure de corrélation spatiale des concentrations en métaux lourds des sols échantillonnés, est estimée par le calcul des semi-variances.

Lorsque l'espacement des points est irrégulier, ce qui est le cas le plus fréquent, on choisit, pour les calculs, un intervalle d'échantillonnage constant, fictif; et on calcule les semi-variances sur la paire de valeurs obtenue (celle de l'échantillon réel et celle calculée de l'échantillon fictif), dans l'intervalle choisi.

4.2.2.4. Les résultats.

* Les mesure de la tendance centrale (distance du point, moyenne, médiane, standard développé, biais, background ou fond) permettent d'étudier les relations de la pollution avec la source principale. Ainsi, les corrélations entre la distance à la source et les concentrations de métaux lourds, montrent une décroissance régulière de la pollution à partir de la source; ici, la source est constituée par les cheminées d'usine. La régularité de cette décroissance est liée au mode de pollution par dépôt atmosphérique.

* Le semi-variogramme des concentrations indique un schéma de dépôt symétrique autour d'un point central d'émission.

* Les valeurs de R permettent d'apprécier la qualité de l'ajustement d'une formule polynomiale à la relation.

* Les valeurs de F donnent le degré de signification statistique de chaque ordre du polynome. Ainsi, on choisit pour degré du polynome, l'ordre à partir duquel une augmentation de degré n'apporte pas d'ammélioration significative.

* La tendance globale de surface donne la significativité de la régression pour chaque degré du polynome; la tendance locale limite cet indice à une fenêtre de dimension inférieure à la distance du semi-variogramme.

Ainsi, pour le barium, les résultats statistiques montrent que la pollution est maximale autour du point d'émission, dont l'activité a été arrêtée en 1975. Une décroissance préférentielle des concentrations suivant la direction NE-SW est due à l'effet des vents dominants.

Le zinc présente un schéma de répartition de base semblable, avec une composante supplémentaire : à l'ouest, une augmentation de concentration due au dépôt de sédiments pollués dans la plaine d'inondation de la Meuse.

4.2.2.5. Conclusion.

Les cartes résultants de ces calculs statistiques, donnent une bonne indication de la distribution générale de la pollution. Mais, elles sont insuffisantes pour déterminer exactement les parcelles répondant aux normes de santé. Par contre, elles constituent une excellente base pour préparer un programme d'échantillonnage détaillé, dans cet objectif de prévision au niveau parcellaire.

4.3. LE TRAITEMENT DE L'IMAGE EN VUE DE LA RESTITUTION GRAPHIQUE.

Le traitement de l'image assure la présentation des données traitées précédemment, sous forme de cartes ou de graphiques, qui seront imprimés, ou consultables à l'écran lors d'interrogation.

Ainsi, les procédures de traitement de l'image comprennent :

- . des procédures de cartographie automatique permettant aux réalisateurs d'élaborer les cartes,
- . des procédures de manipulation des images accessibles aux interrogateurs, leur permettant de choisir la représentation graphique des données la mieux adaptée à leur problème.

Parmi les manipulations d'image possibles, figurent :

- . les changements d'échelle,
- . les changements de style de projection cartographique (projection conforme Lambert conservant les angles, ou projection équivalente d'Abert conservant les surfaces),
- . la visualisation des fonds de carte,
- . les regroupements d'images,
- . le zoom,
- . le changement de couleurs et de trames des images,
- . le passage d'une représentation à 2 dimensions (coupes, cartes) à une représentation à 3 dimensions (blocs-diagrammes).

Le module de cartographie automatique du système SAS, déjà cité, offre de nombreuses fonctions de ce type.

* Un module GRAPH assure la représentation graphique des séries statistiques et de leurs diagrammes spécifiques.

* 3 procédures assurent la préparation des données graphiques :

- . GPROJECT permet de choisir le système de projection cartographiques de données graphiques sphériques;
- . GREMOVE réalise l'agrégation des unités spatiales selon un code définissant le niveau d'emboîtement choisi, puis efface les contours résiduels;

. GREDUCE est une procédure de généralisation des contours;

* 3 procédures centrales assurent la réalisation des représentations graphiques :

- . G3 trace des surfaces à 3 dimensions;
- . GCONTOUR visualise les surfaces dans un plan par des courbes de niveau. Il permet de passer à des blocs-diagrammes, qu'il est possible de faire pivoter autour de n'importe lequel des 3 axes. On peut ainsi représenter la piézométrie d'une nappe, la lame d'eau des précipitations, etc...

Si l'échantillonnage n'est pas régulier, la cote z sera estimée par un calcul statistique de type régression, soit à partir des 30 points d'échantillonnage les plus proches du point à estimer, soit à partir d'une équation de surface de tendance;

- . GMAP assure le tracé à partir des variables à représenter, qui sont stockées dans un tableau permanent, et du fond numérisé stocké dans un autre tableau; ce module offre à choisir :

- + le mode de représentation : CHORO, PRISM, BLOCK, SURFACE,
- + la manière de découper en classes les valeurs des variables à représenter,
- + le nom à attribuer aux variables à représenter;

* Après ces 3 procédures centrales, viennent 3 procédures permettant l'habillage des représentations graphiques :

- . PATTERN est appelé pour chaque classe à représenter, il propose :
 - + un choix de 8 couleurs,
 - + un choix de trames (croisillons, hâchures, vides, pleins);
- . TITLE règle la présentation du titre;
- . FOOTNOTE réalise la légende.

Le traitement de l'image comprend également des calculs graphiques sur écran, tels que les calculs de distance, de périmètre, de surface.

Dans la version 5 du système SAS, le traitement de l'image s'applique également à la gestion d'un catalogue d'images cartographiques. L'utilisateur dispose des possibilités suivantes :

- . affichage de tout graphique du catalogue;
- . affichage d'un groupement de plusieurs graphiques sur la même page, avec le patron d'assemblage;
- . groupement de plusieurs pages suivant un affichage séquentiel, tel que le "fondu enchaîné" des diapositives;
- . habillage des cartes avec identification de chaque point.

Les images obtenues par ces différentes manipulations sont virtuelles, et indépendantes du matériel de sortie, que ce soit un écran ou un traceur. Le passage de ces images virtuelles aux images concrètes obtenues à l'écran ou sur l'imprimante est réalisé par des pilotes d'unités graphiques; ce sont des

programmes interprétant le tracé virtuel pour les différents organes de sortie.

5. LES ORGANES DE SORTIE ET LES MODES DE RESTITUTION.

La sortie des résultats peut s'effectuer sur écran graphique pour visualisation, ou donner lieu à une édition imprimée sur papier, à une restitution sur film ou sur support magnétique sous forme de fichiers (M. Souris, 1986).

Les sorties sur écran graphique de type Péricolor ou Techtronix, nécessitent un moniteur vidéo de haute définition, des mémoires graphiques à affichage rapide avec un grand nombre de couleurs, et un logiciel spécialisé. L'écran graphique utilisé par le système TIGER est un GIXI radiance 320 à 256 couleurs.

Les imprimantes et les traceurs sur support papier nécessitent un système de copie d'écran (hard copy). Les imprimantes graphiques peuvent être à aiguilles, à marteau, à jet d'encre. Les traceurs à plat ou à rouleau permettent tous les types d'implantation : ponctuelle, linéaire ou zonale.

Un autre type de sortie peut être envisagé : le télédéchargement vers des microprocesseurs de terrain, reliés à l'ordinateur central par une ligne télécom. La version 6 du système SAS déjà cité, assure cette fonction avec des microordinateurs IBM PC XT/370 ou AT/370, à partir d'un système central sous VM / CMS .

6. LES ORIENTATIONS FUTURES DES SYSTEMES D'INFORMATION GEOGRAPHIQUE.

A travers les étapes précédentes, menant de la récolte des données brutes à l'obtention de cartes de vulnérabilité, sont apparues les facilités et les possibilités nouvelles offertes par l'infographie, ou cartographie assistée par ordinateur.

6.1. LES AVANTAGES DIVERS OFFERTS PAR L'INFOGRAPHIE.

Le volume des informations stockées est considérable, et ces informations sont immédiatement disponibles, consultables et utilisables.

La gestion des informations stockées (recherche et mise à jour, insertion, mise en relation) est rapide et facile. Les possibilités de sélection de l'information sont multiples. Les divers programmes de calculs statistiques et de modélisation sont applicables à n'importe quelles variables statistiques, à tout moment de la consultation. L'utilisateur a la possibilité de choisir lui-même les types de représentation graphique qui l'intéressent. Il peut en faire varier toutes les caractéristiques pour sélectionner les représentations graphiques qui lui conviennent le mieux.

6.2. LES CROISEMENTS DE CARTES THEMATIQUES.

L'intégration d'images numériques en provenance d'une autre base et la mise en relation des données localisées de diverse nature, ouvrent la voie au croisement de cartes thématiques de diverses disciplines, telles que :

- . les cartes météorologiques et les cartes agricoles,
- . les cartes de pollution des eaux de surface et les cartes de vulnérabilité des eaux souterraines.

Certains organismes se sont penchés plus spécialement sur le croisement des cartes thématiques. Il s'agit, entre autres, de :

- . l'INRA (Institut national de recherche agricole),
- . le Centre de Télédétection et d'Analyse du Milieu Naturel,
- . l'Ecole Nationale Supérieure des Mines de Valbonne.

Dans le domaine de la vulnérabilité des nappes, 3 modèles ont été réalisés à partir de croisement de cartes (A. Breewsma et W. Van Duyvenbooden, 1987) :

- . un modèle régional de transport des phosphates,
- . un modèle régional d'infiltration des nitrates,
- . un modèle régional d'acidification des sols.

Les 2 premiers modèles ont été obtenus en croisant :

- . des cartes de vulnérabilité des eaux souterraines par infiltration, basées sur la capacité d'échange des cations du sol et sur le temps de transit des eaux d'infiltration,
- . avec des cartes de précipitations,

- . avec des cartes d'utilisation des sols,
- . avec des cartes caractérisant les entrées des polluants dans le sol.

Le troisième modèle est en relation avec une carte de risque d'acidification des aquifères (M. Holmberg et coll., 1987).

Cette dernière a été obtenue en croisant une carte de sensibilité à l'acidification des aquifères avec une carte des précipitations acides. La carte de sensibilité a été, elle-même, élaborée à partir des données suivantes :

- . la carte des sols de la FAO, digitalisée,
- . la carte hydrogéologique internationale de l'Europe, digitalisée.

Ainsi, le croisement de cartes thématiques et la manipulation des informations qu'elles contiennent, apparaissent comme des techniques essentielles pour l'étude de la vulnérabilité des eaux souterraines aux polluants. L'une des finalités de l'étude de vulnérabilité est la modélisation de la qualité des eaux souterraines pour une période donnée.

6.3. LES SYSTEMES EXPERTS.

Une nouvelle génération de système informatique a fait son apparition, le système expert, radicalement différent des précédents par son fonctionnement.

Il s'agit d'un système convivial, reposant sur un échange continu de questions posées par l'ordinateur et de réponses fournies par l'utilisateur. Ces réponses de l'utilisateur, constituent les données que le système traite au fur et à mesure de l'échange, et dont dépendent les questions posées ultérieurement par le système.

Au terme de la série des questions système - réponses utilisateur, le système rend un verdict sur le problème défini par cet ensemble de questions-réponses. Ce verdict est fiable, si le problème posé figure parmi les cas ou scénarios gérés par le système expert.

Les applications des systèmes experts dans le domaine de la vulnérabilité des eaux souterraines aux polluants sont encore très rares.

W. Mak et A.P. Bot (1987) annonce un système expert d'évaluation de la vulnérabilité des eaux souterraines. Les données sollicitées portent sur le niveau piézométrique, le type de sol, le schéma de drainage, l'utilisation des terres, la végétation, l'altitude relative, et la pente. A partir de ces données, le système fournit les caractéristiques de la migration verticale et de la migration horizontale des eaux polluées. La nature des polluants ne semble pas prise en compte, de même que la structure des terrains alentour. Cela laisserait supposer que le modèle correspond à un aquifère homogène, de structure régulière, et que les polluants sont entièrement solubles dans l'eau.

Un système de même type, a été produit par W.A.M. Kerkhof et coll. (1987) de la même équipe de recherche que les précédents auteurs. Il concerne l'évaluation des effets de dépôts d'ordures sur l'environnement au niveau du site d'une

décharge. Ce système a été bâti à partir de l'analyse de 52 sites de décharges à travers la Belgique.

Les questions posées concernent :

- . la production de polluants sur le site; entre autres, la nature des déchets déposés (ces données déterminent 5 facteurs);

- . les trajets possibles des polluants; entre autres :
 - + les dimensions du site,
 - + la profondeur de l'eau,
 - + les dimensions et l'intensité du système de

drainage,

hydrauliques,

- + le système hydrologique et les potentiels

- + l'épaisseur et la nature de la couche de

couverture des déchets;

ces données sur les trajets de polluants déterminent 14 facteurs;

- . l'utilisation du site et des alentours; entre autres :

- + la végétation ("A-t-elle subit des dégâts visibles, par les gaz?"),

- + l'utilisation des sols,

- + la gestion des eaux de surface,

- + l'utilisation des eaux souterraines;

ces données sur l'utilisation du site déterminent 6 facteurs.

Le système fournit ensuite, une évaluation des risques pour les gaz, le sol, les eaux de surface, les eaux de la zone d'infiltration, le premier et le deuxième aquifères.

Les systèmes experts présentent un mode de fonctionnement très souple et très convivial. Son avantage prépondérant réside dans le fait qu'il ne nécessite de la part de l'utilisateur aucune connaissance des mécanismes de propagation des polluants; il suffit à l'utilisateur de pouvoir répondre aux questions.

En revanche, si le problème soumis au système ne figure pas précisément parmi les cas de figure envisagés par le système, la réponse fournie risque d'être totalement fautive, sans pour autant que le système ne l'assortisse d'un message de mise en garde. Les limites du domaine dans lequel le système expert est habilité à se prononcer, doivent être mises en exergue par les concepteurs du système, afin de prévenir tout malentendu pour les usagers.

CONCLUSION

Cette note de synthèse a suivi les différentes étapes de conception et de fonctionnement d'un système adapté à la cartographie assistée par ordinateur.

Les contraintes sur l'aspect spatial des données et sur les organes de gestion de ces données, ont été détaillées; car, ce sont eux qui permettront les croisements de cartes thématiques. L'établissement des relations entre les données descriptives, est un travail d'analyse qui revient à l'hydrogéologue, et qui conditionne l'étendue des possibilités du système.

Les travaux présentés au colloque de l'ORSTOM donnent un aperçu global et complet du fonctionnement d'un tel système. Les travaux de l'Institut fuer Angewandte Geodasie de Frankfort fournissent de multiples exemples d'applications. Parmi celles-ci, la réalisation de l'Atlas de la Suisse illustre de façon détaillée le fonctionnement du système de gestion des données dans la production des différentes cartes thématiques de l'Atlas.

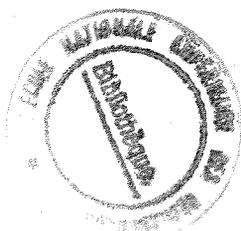
Les travaux sur la pollution des eaux souterraines ont permis de dégager les paramètres les plus importants qui interviennent dans les phénomènes de propagation ou d'élimination des polluants.

Les relations qui lient certains de ces paramètres apparaissent dans les travaux de modélisation des processus physico-chimiques et des écoulements. Ces travaux ne permettent pas de dégager un consensus sur la représentation de ces phénomènes.

Certaines études simplifient les problèmes, en ne s'intéressant qu'au devenir du polluant soluble en phase aqueuse. Cette simplification est inacceptable pour traiter les pollutions par les PCB, les détergents et les hydrocarbures. Ces polluants, en se partageant, engendrent des phases huileuses non miscibles.

Les travaux qui s'intéressent au devenir de ces phases, s'appuient sur des études de terrain, souvent longues et coûteuses, et ne proposent pas de modèle généralisable. Les travaux de laboratoire sur ces types de pollution propose des modèles qui sont soumis à caution, par le fait que l'hétérogénéité des terrains rend souvent impossible leur application directe aux cas concrets.

Les études de pollution par les PCB, les détergents et les hydrocarbures passent par des études de terrain ponctuelles, dont les résultats doivent être traités par les méthodes géostatistiques pour arriver à un modèle. Celui-ci pourra alors servir de base à des représentations cartographiques, ou à établir la base de faits d'un système expert.



BIBLIOGRAPHIE

- BRASSEL, K. ; DORIGO, G. ; HERZOG, A. - 1986
Computer-Assisted Construction of a Structural Atlas of Switzerland.
In : Contribution to Geodesy, Photogrammetry, and Cartography ; Series 2 ; no. 44 ; p. 15-34.- ISSN 0469-4244.
EURO CARTO IV Seminar ; Frankfurt am Main, West Germany ; Oct 15-18, 1985.
- GELHAR, L.W. ; GUTJAHR, A.L. ; NAFF, R.L. - 1979
Stochastic analysis of macrodispersion in a stratified aquifer.
Water Resources Research ; 1979 ; 15 ; 6 . - ISSN 0043-1397.
- GVIRTZMAN, H. ; RONEN, D. ; MAGARITZ, M. - 1986
Anion exclusion during transport through the unsaturated zone.
Journal of Hydrology ; 1986 ; 87 ; 314 ; p. 267-284.
- JACKSON, R.E. (ed.) - 1980
Aquifer contamination and protection.
In : Studies and reports in hydrology ; 30 .- UNESCO ; Paris .- 440 p.
- MATHERON, G. ; MARSILLY, G. de - 1980
Is transport in porous media always diffusive ? A counter exemple.
Water resources Research ; 1980 ; 16 ; 4 . - ISSN 0043-1397
- SOURIS, M. - 1986
Système d'information géographique et bases de données.
In : Traitement des données localisées : l'infographie à l'ORSTOM. - Colloque ORSTOM ; Paris ; 1986 .
- WELBY, C.W. - 1987
Use of ground water quality monitoring data in determination of decay coefficients for organic chemicals in shallow ground water.
In : Vulnerability of soil and groundwater to pollutants, Noordwijk aan Zee, Netherlands ; March 30-April 3, 1987. - p.861-868.



BIBLIOTHEQUE DE L'ENSSIB



966021D