

Module de Doctorat 3^{ème} Cycle LMD

Les systèmes d'information géographiques comme des outils d'analyse, de représentation et de modélisation pour les prises de décisions

Présenté par : Prof. Krolladi Mohamed-Khireddine
Département d'Informatique
Facultés des Sciences Exactes
Université Echahid Hamma Lakhdar d'El Oued
Tél. 0770314924
Email. krolladi@univ-eloued.dz et krolladi@yahoo.fr
Site Web. www.univ-eloued.dz
<http://krolladi.doomby.com/> et <http://krolladi.e-monsite.com/>



Les systèmes d'information géographique comme des outils d'analyse, de représentation et de modélisation pour les prises de décisions

I – Introduction

- I.1 – De la cartographie aux systèmes d'information géographique
- I.2 – Les six A
- I.3 – De la CAO (conception assistée par ordinateur) au SIG (système d'information géographique)
- I.4 - Les données structurées en couches thématiques dans les systèmes d'information géographique

II – Qu'est-ce qu'un système d'information géographique ?

- II.1 – Quelques définitions

III – Potentialités des systèmes d'information géographique

- III.1 – Une analyse à des échelles pertinentes
- III.2 - Une compréhension spatiale des données
- III.3 - Une représentation graphique signifiante

IV - Limites des systèmes d'information géographique

- IV.1 - Les problèmes liés au fond de plan
- IV.2 - Les problèmes de disponibilité ou de fiabilité des données
- IV.3 - Les problèmes liés à la représentation graphique
- IV.4 - Les problèmes reliés aux systèmes d'information géographique

V - Utilité des systèmes d'information géographique dans la constitution d'un diagnostic de territoire

- V.1 - Croiser des données
 - V.2 - Représenter certains résultats
 - VI - La place des systèmes d'information géographique
 - VI.1 - Géomatique
 - VI.2 - SIRS, SIT et SIG
 - VII- L'originalité des systèmes d'information géographique
 - VIII - Les domaines d'application des systèmes d'information géographique
 - IX - La représentation de la réalité dans un système d'information géographique
 - IX.1 – Les systèmes d'information géographique comme outils dans la démarche géographique
 - X - Les systèmes d'information géographique comme langage
 - X.1 - L'information géographique
 - X.2 - Le lien entre signifiant et signifié
 - XI - Les niveaux de la modélisation géographique
 - XI.1 - Les niveaux de la modélisation géographique
 - XI.2 - Les modèles descriptifs
 - XI.3 - Les modèles explicatifs
 - XI.4 - Les modèles prédictifs
 - XII - Un outil d'action pour l'aménagement d'un territoire
 - XII.1 - Les logiques de l'aménagement du territoire
 - XII.2 - Aspects spatiaux : concepts spatiaux
 - XII.3 - Découpages irréguliers
 - XII.4 - Découpage régulier : l'approche par maille
 - XII.5 - Relations entre classes d'objets spatiaux
 - XII.6 - Un autre modèle de données : le TIN
 - XII.7 - Aspects sémantiques : les attributs
 - XII.8 – Démarche de construction de la base de données GEOSPATIALE
 - XII.9 - Relations spatiales entre les objets
 - XII.10 - Entité spatiales selon l'approche topologique et non topologique
 - XIII – Cartomatique et sémiologie graphique
 - XIII.1 – Notions de base
 - XIII.2 – Contenu de la présentation
 - XIII.3 - Objectifs et Qualités d'un SIRS (système d'information à référence spatiale)
 - XIII.4 - Quelques domaines d'application des SIRS
 - XIII.5 - Inventaire, analyse et communication
-

- XIII.6 - La sémiologie graphique : une grammaire du langage graphique
- XIII.7 - Les objets cartographiques
- XIII.8 - Une typologie des objets en mode vectoriel
- XIII.9 - Les données thématiques
- XIII.10 - Les métadonnées
- XIV – Les phases de création d'une carte thématique
 - XIV.1 - Traitement des données thématiques
 - XIV.2 - Traitement de la géobase
 - XIV.3 - Discrétisation et représentation cartographique
 - XIV.4 - Le processus cartographique dans MapInfo
 - XIV.5 - Un exemple dans MapInfo La population d'un centre urbain en 1991
- XV – Généralités sur les bases de données géographiques
 - XV.1 - Architecture des bases de données géographiques
 - XV.2 - La base de données géographique : l'analyse
- XVI – Les systèmes de projections et les référentiels

I – Introduction

La connaissance du territoire communal peut se scinder en deux types :

1. La connaissance générale du fonctionnement du territoire, de type informel ou qualitatif, et non localisée ;
2. Et la connaissance technique, de type statistique ou géographique, et localisée.

I.1 – De la cartographie aux systèmes d’information géographique

Les travaux sur les cartes datent depuis pratiquement les années 50, c’est-à-dire depuis l’avènement de l’informatique. Mais, on commencera notre histoire à l’aube des systèmes d’information géographique selon le tableau ci-dessous.

Périodes	Activités
1980	Numérisation de l’information (de la carte papier vers la carte numérique grâce aux outils de CAO et de DAO)
1990	Intégration des informations numériques : les bases de données géographiques (SGBD Géo) et les informations cartographiques numérisées (CAO ou DAO) pour la création des systèmes d’information géographique (SIG)
2000-2010	Passage des informations de logiciels de bureau vers les services Web pour créer les bases de données géographiques (BD géo) et les portails géographiques.
2011 -2020	Intégration des informations des villes dans le cadre des Smart cyties (les villes connectées) pour la gestion moderne de la ville.

I.2 – Les six A

On peut résumer le passage de cartographie aux systèmes d’information géographique selon le tableau ci-dessous.

Les six A	Activités des systèmes d’information géographique
Acquisition	Saisie des informations des informations géographiques sous forme numérique
Archivage	Gestion de base de données
Analyse	Manipulation et interrogation des données géographiques. Cela revient à répondre aux questions suivantes : où ? quoi ? comment ? quand. Et si ?

Affichage	Mise en forme et visualisation
Abstraction	Représentation du monde réel
Anticipation	Prospective

I.3 – De la CAO (conception assistée par ordinateur) au SIG (système d’information géographique)

Les questions de base auxquelles un SIG doit répondre

1. Quoi ? A tel ou tel endroit, que trouve-t-on ? Inventaire de tous les objets de l’espace concerné, nature des objets selon leur localisation, quantité, superpositions, proximités.
2. Comment ? Quelles relations existent entre les objets ou phénomènes ? Que traduit leur répartition spatiale ? Problématique de l’analyse spatiale, au cœur de la Géographie.
3. Quand ? Des changements sont-ils intervenus, et si oui, quand ? Âge, évolution de tel objet ou phénomène, historique et actualisation des données, analyse temporelle.
4. Où ? Inventaire d’un type d’objet dans tous les endroits où il se trouve, mise en évidence de sa répartition spatiale.
5. Et si ? Que se passerait-il si tel scénario d’évolution se produisait ? Projection dans l’avenir, simulation, études de projet ou d’impact.

Les deux premières questions ressortent «mécaniquement» de la cartographie. La troisième question est essentielle (mise en évidence et explicitation éventuelle de corrélations). La quatrième question pose des problèmes de représentation. La cinquième question est délicate à traiter (problème de la simplification, apparition d’événements contingents...). Les informations traitées dans un système d’information géographique sont généralement disposées en couches homogènes, rassemblant un même type d’objet.

I.4 - Les données structurées en couches thématiques dans les systèmes d’information géographique

L’information géographique est constituée de couches d’information représentant les multiples faces de l’environnement social et physique, c’est l’image du millefeuille comme sur la figure I.1.

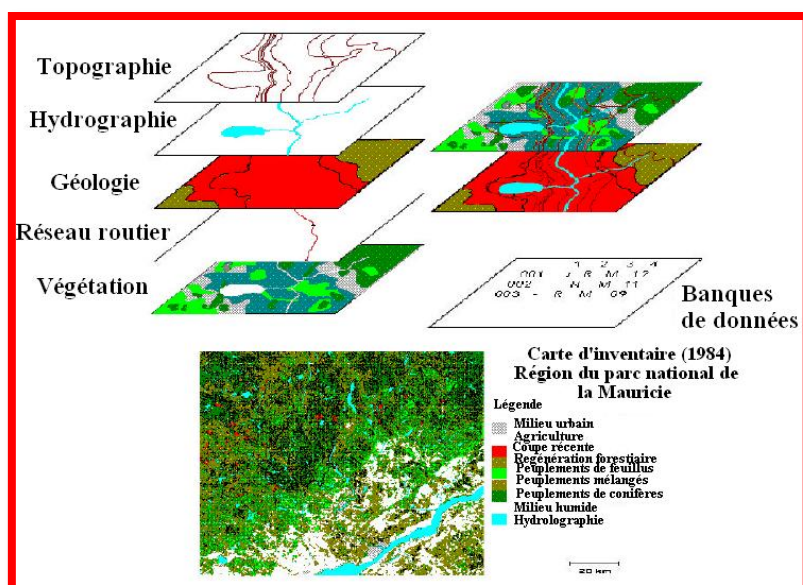


Figure 1.1 – Les couches thématiques d'un système d'information géographique

II - Qu'est-ce qu'un système d'information géographique ?

La connaissance technique du territoire peut être appréhendée à l'aide d'un Système d'Information Géographique (SIG). Il peut permettre, concernant des données, de répondre aux questions : où, quoi, comment, quand, et si ? Un système d'information géographique est un ensemble de données, d'outils et de moyens permettant de stocker, gérer, croiser, représenter une information localisée sur le territoire.

Il fait appel à :

1. Un référentiel géographique, selon que l'on s'intéresse au parcellaire, aux îlots, aux routes, aux cours d'eau, etc. Les plus courants sont le RGE (référentiel à grande échelle) ou la BD Carto de l'INC, les IRIS 200 ou BaseIlot de l'INSEE, voire le Cadastre numérisé (PCI Vecteur de la DGI, BD Parcellaire de l'INC).
2. Des données ou phénomènes localisables, pouvant être décrits par leur nature, leur aspect et leurs caractéristiques
3. Et des outils de gestion / calcul / affichage (logiciels).

Son utilisation suppose donc de mobiliser différents métiers, tels que la saisie des données, la gestion des données, l'analyse des données (techniciens – géographes ou statisticiens) et la représentation des données ou des résultats (graphistes).

L'acronyme SIG signifie "Système d'Information Géographique". C'est la traduction de l'acronyme anglais GIS, qui signifie à la fois Geographic Information Systems et Geographic Information Science.

II.1 – Quelques définitions

"Système d'information : système de communication permettant de communiquer et de traiter l'information" selon la norme internationale ISO 5127-1-1983.

"Information géographique : L'information géographique est la représentation d'un objet ou d'un phénomène réel, localisé dans l'espace à un moment donné" selon la cartographie numérique et information géographique en 1994.

"Géomatique : Ensemble des techniques de traitement informatique des données géographiques" selon le JO du 14 février 1994.

"Système d'informations géographiques : Système informatique de matériels, de logiciels et de processus conçus pour permettre la collecte, la gestion, la manipulation, l'analyse, la modélisation et l'affichage de données à référence spatiale afin de résoudre des problèmes complexes d'aménagement et de gestion" selon le comité fédéral de coordination inter agences pour la cartographie 1988.

"Un ensemble de procédures utilisées pour conserver et traiter de l'information à référence géographique" (selon Aronoff, 1989).

"Un ensemble puissant d'outils pour saisir, conserver, extraire, transmettre et afficher les données spatiales décrivant le monde réel" (selon Burrough, 1986).

"Une structure institutionnelle qui utilise la technologie des systèmes d'information géographique pour gérer une banque de données et exercer son mandat spécifique avec un financement stable et continu" (selon Carter, 1989).

"Les systèmes d'information géographique comportent trois types de composantes technologiques (matériel et logiciel), informatives (bases de données géographiques et associées) et infrastructurelles (personnel, installations, services de support) » (selon Dickinson et Calkins, 1988).

"Un type particulier de système d'information dont la base de données contient des informations reliées à des entités physiques, des activités ou des événements localisés et assimilables aux formes géométriques de points, de lignes et de zones. Un système

d'information géographique gère les informations spécifiques à ces points, lignes et zones pour extraire les données requises afin de réaliser des recherches et des analyses spécialisées" (selon Dueker, 1979).

"Un système doté de fonctions de modélisation spatiale puissantes" (selon Koshkariov et al., 1989).

"Un ensemble de procédures informatisées qui offre aux professionnels des fonctions adaptées pour le stockage, l'extraction, le traitement et l'affichage des données à référence spatiale" (selon Ozemoy et al., 1981).

"SIG : Ensemble de données repérées dans l'espace, structuré de façon à pouvoir en extraire commodément des synthèses utiles à la décision" selon les études du CNIG en 1990.

La vocation d'un système d'information géographique est de rassembler au sein d'un outil informatique des données diverses, localisées dans le même espace géographiques, relatives à la Terre et à l'homme, à leurs interactions et à leurs évolutions respectives. La finalité d'un SIG est l'aide à la décision. On peut dire en résumé qu'un système d'information géographique peut désigner un ensemble complet de composantes permettant les traitements informatiques de l'information géoréférencée, comme sur la figure II.1).

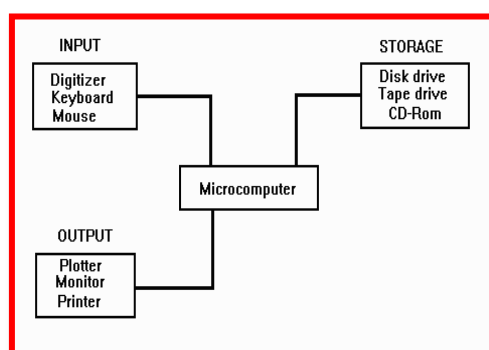


Figure II.1 – Les composantes d'un système d'information géographique

III - Potentialités des systèmes d'information géographique

III.1 - Une analyse à des échelles pertinentes

Les systèmes d'information géographique permettent de travailler à une échelle supra-communale, qui est souvent le niveau de disponibilité et de validité de nombreuses données. Ceci sera utile pour certaines approches thématiques dont les phénomènes ne s'arrêtent pas aux contours administratifs d'une commune, comme par exemples : la mise en évidence de

corridors biologiques, les croisements de données géologiques et d'occupation des sols et les approches des pollutions ou nuisances (bruit, air, risques, etc.). Le choix du référentiel sera donc lié à l'échelle d'analyse pertinente des phénomènes observés.

III.2 - Une compréhension spatiale des données

Il est ainsi possible d'utiliser les systèmes d'information géographique pour croiser diverses sources d'information, les analyser et représenter des phénomènes complexes. L'utilisation des systèmes d'information géographique permet une compréhension plus facile de certaines données souvent abstraites, par leur représentation graphique sur un référentiel géographique. D'autant plus que le fond de plan utilisé pourra être intelligemment enrichi par la localisation des cours d'eau, équipements, bâtiments publics, routes, etc. éléments permettant un repérage aisé dans l'espace par le destinataire ou l'utilisateur de la carte.

III.3 - Une représentation graphique signifiante

Selon les logiciels utilisés, les systèmes d'information géographique permettront de représenter de nombreux objets : polygones et surfaces, lignes, points, texte, etc. Il sera fondamental de porter attention à la sémiologie graphique retenue. On pourra jouer sur les couleurs, les épaisseurs, les textures des différents objets. On pourra également utiliser des représentations symboliques (flèches, patates, etc.). Ainsi la représentation des résultats au travers des systèmes d'information géographique peut donner quelque chose de très signifiant, à condition de faire intervenir une compétence adéquate, de graphiste par exemple.

IV - Limites des systèmes d'information géographique

IV.1 - Les problèmes liés au fond de plan

Ils peuvent être de diverses sortes : qualité, précision et disponibilité des fonds de plan échelle (et adaptation à l'affichage des données représentées à cette échelle), continuité territoriale (par exemple pour le cadastre entre communes) ou mise à jour des fonds de plan (ex. du cadastre), qui peut entraîner des coûts et des besoins éventuels en compétences pour leur gestion.

IV.2 - Les problèmes de disponibilité ou de fiabilité des données

La fourniture des données nécessaires à l'observation ou l'analyse effectuée grâce aux systèmes d'information géographique peut être problématique, en fonction de la disponibilité

des données, de la discontinuité territoriale, du coût d'obtention, de l'existence de bases alimentées régulièrement, et de la qualité de ces données origine, fiabilité du fournisseur, mise à jour, précision, exhaustivité, etc.

IV.3 - Les problèmes liés à la représentation graphique

Les logiciels utilisés comme les choix de modes de travail peuvent avoir diverses conséquences sur la représentation parfois limitée, en fonction des logiciels utilisés trames, dégradés, schématisation, etc., entraînant le recours à l'utilisation de logiciels purement graphiques et des difficultés de communication, liés au choix par exemple de la couleur ou du noir et blanc avec les problèmes d'impression correspondants.

IV.4 - Les problèmes reliés aux systèmes d'information géographique

Ils sont de plusieurs ordres :

- Normes (ISO, ACNOR, ANSI, IEEE, OSF)
 - Systèmes d'exploitation
 - Interface usager
 - Réseaux
 - Langages (SQL)
 - Formats d'échanges des cartes numériques
- Terminologie
 - Pose des défis considérables en raison de la vitesse des développements technologiques
 - Termes impropres
 - Vocabulaire immature
- Politiques
 - Accès aux données (diffusion, coûts)
 - Contrôle de l'information et confidentialité
 - Formats d'échanges privilégiés
- Légaux
 - Protection de la vie privée
 - Droits d'auteur
 - Responsabilité civile

V - Utilité des systèmes d'information géographique dans la constitution d'un diagnostic de territoire

V.1 - Croiser des données

L'utilisation des systèmes d'information géographique peut être un moyen intéressant de croiser des résultats ou données relevant de thématiques voire de périodes différentes. Elle permet ainsi de mettre en évidence des évolutions, de relier des phénomènes, voire de les expliquer. Les tableaux ou grilles thématiques présentés dans le cédérom joint, mettent en évidence des liens ou interactions fréquentes entre thèmes. Ils fournissent également des références et indiquent des structures gestionnaires de données qui pourront alimenter des systèmes d'information géographique. En outre, les données locales, comme les données nationales (BD Carto, BD Topo, CORINE Land Cover, etc.) pourront fournir des informations pour répondre aux questions posées dans les grilles.

V.2 - Représenter certains résultats

Même si parfois analyse et représentation peuvent être faites à l'aide d'un même logiciel (MapInfo, InfoMap, Arcview, Géoconcept, etc.), l'analyse se fait souvent avec un logiciel spécifique système d'information géographique, et la représentation est réalisée avec d'autres outils (PAO, DAO, AutoCad, etc.). Notons que cette dernière nécessite souvent une compétence de graphiste, de manière à transformer le résultat de l'analyse et le message à faire passer, en représentation signifiante. Dans la présentation des résultats, on portera donc une attention particulière à la fois à la sémiologie graphique, à la légende des plans et à la mise en page des cartes.

VI - La place des systèmes d'information géographique

VI.1 - Géomatique

L'avènement de la géomatique a suivi celui de l'informatique. La géomatique a d'abord consisté à l'automatisation de la cartographie. Par la suite, les progrès informatiques ont permis d'intégrer une masse croissante de données localisées, rendant nécessaire l'usage de Bases de Données, ouvrant la voie à la conception de SIG. C'est une discipline ayant pour objet la gestion des données à référence spatiale par l'intégration des sciences et des technologies reliés à leur acquisition, leur stockage, leur traitement et leur diffusion. Elle intègre principalement les disciplines telles que les mathématiques, la physique,

l'informatique, la topométrie, la cartographie, la géodésie, la photogrammétrie et la télédétection selon Bergeron en 1992.

VI.2 - SIRS, SIT et SIG

a - SIRS

"Un système d'aide à la décision qui intègre des données à référence spatiale dans un environnement informatique pour la solution de problèmes spécifiques" (selon Cowen, 1988).

b - SIT

"Un instrument de décision dans les domaines juridique, administratif et économique. Il est généralement axé sur une gestion vectorielle des données cadastrales" (selon Bédard, 1987).

c - SIG

"Un ensemble de principes, de méthodes, d'instruments et de données à référence spatiale utiliser pour saisir, conserver, transformer, analyser, modéliser, simuler et cartographier les phénomènes et les processus distribués dans l'espace géographique. Les données sont analysées afin de produire l'information nécessaire pour aider les décideurs" (selon Thériault, 1992).

VII- L'originalité des systèmes d'information géographique

On va voir par rapport à la géomatique, à la cartomatique et à la télédétection :

- Les objectifs : Les systèmes d'information géographique ont comme mission fondamentale de traiter les données géographiques afin de construire des modèles utilisables pour effectuer des simulations et assister les décideurs.
- Les moyens : Les systèmes d'information géographique intègrent obligatoirement plusieurs outils d'analyse spatiale afin d'effectuer la synthèse des données et dégager les traits dominants de l'organisation des paysages terrestres.

VIII - Les domaines d'application des systèmes d'information géographique

Ils sont aussi nombreux que variés et il serait impossible d'en dresser ici une liste exhaustive. Citons cependant :

- Tourisme (gestion des infrastructures, itinéraires touristiques)
- Marketing (localisation des clientèles)
- Milieu scolaire (parcours des autobus scolaires)
- Inventaire et gestion des ressources (équipements urbains)
- Planification urbaine (zonage)
- Gestion des installations (AM/FM, Automatic Mapping and Facilities Management)
- Protection civile (gestion et prévention des catastrophes)
- Politique (cartes électorales)
- Transport (planification, gestion des réseaux et conduite assistée)
- Hydrologie (gestion des réserves hydroélectriques)
- Réseaux de distribution (gaz, électricité, marchandises)
- Foresterie (gestion des coupes et sylviculture)
- Géologie (prospection minière)
- Biologie (études du déplacement des populations animales)
- Défense (planification des opérations militaires)
- Etc.

IX - La représentation de la réalité dans un système d'information géographique

"L'information géographique est la représentation d'un objet ou d'un phénomène réel, localisé dans l'espace à un moment donné" selon la cartographie numérique et information géographique en 1994. L'information géographique est un ensemble reliant une information relative à un objet, décrit par sa nature, son aspect, ses relations éventuelles avec d'autres objets ou phénomènes. C'est le niveau sémantique. L'ensemble des descripteurs de l'objet forme ses attributs. La forme et la localisation de l'objet sur la surface terrestre, exprimés dans un systèmes de coordonnées explicite. C'est le niveau géométrique. Les définitions « officielles » intègrent toujours ces deux niveaux (sémantique et géométrique). L'information géographique peut être définie comme une représentation de la réalité localisée dans le temps et dans l'espace comme sur la figure IX.1.



Figure IX.1 – La représentation de la réalité dans un système d’information géographique

IX.1 – Les systèmes d’information géographique comme outils dans la démarche géographique

En tant qu’outils, les systèmes d’information géographique font partie de la panoplie à disposition du géographe pour la connaissance du territoire. Elle a comme finalité une action sur le territoire. Il est intéressant de situer les systèmes d’information géographique dans une démarche géographique qui implique au moins trois niveaux de connaissance comme IX.2 :

- Le niveau conceptuel : explicitation des concepts et théories (issus des sciences naturelles et sociales) utilisés dans la connaissance du territoire, et de la forêt tropicale comme cas particulier.
- Le niveau empirique : il concerne les moyens de mesure et d’accumulation de données sur le territoire.
- Et le niveau de modélisation et simulation : une fois les concepts explicités et les données à disposition, la modélisation des systèmes d’information géographique peut commencer.

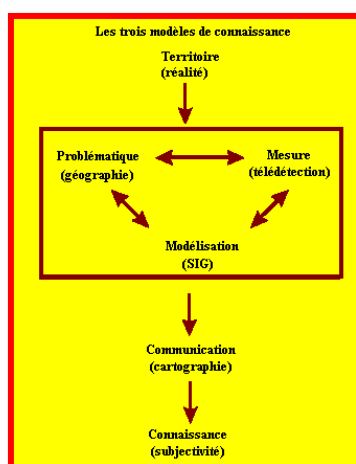


Figure IX.2 – La démarche géographique en trois niveaux

X - Les systèmes d'information géographique comme langage

Les systèmes d'information géographique, en tant que langage, permettent la communication qui est généralement définie comme la transmission de messages d'un émetteur vers un récepteur à travers un canal de transmission et à l'aide d'un code. Le problème de la communication est triple:

- Niveau 1 : Avec quelle exactitude les symboles de la communication peuvent-ils être transmis? Ce qui implique un problème technique.
- Niveau 2 : Avec quelle précision les symboles transmis véhiculent-ils la signification désirée? Ce qui implique un problème sémantique.
- Et niveau 3 : avec quelle efficacité la signification reçue influence-t-elle la conduite dans le sens désiré? Ce qui implique un problème de l'efficacité.

X.1 - L'information géographique

Dans le cas de l'information géographique, le signal est constitué par la donnée numérique (par ex. une série de coordonnées X, Y et Z) et le sens par le concept géographique qu'il est censé véhiculer, par ex. ces coordonnées représentent une forêt comme sur la figure X.1.

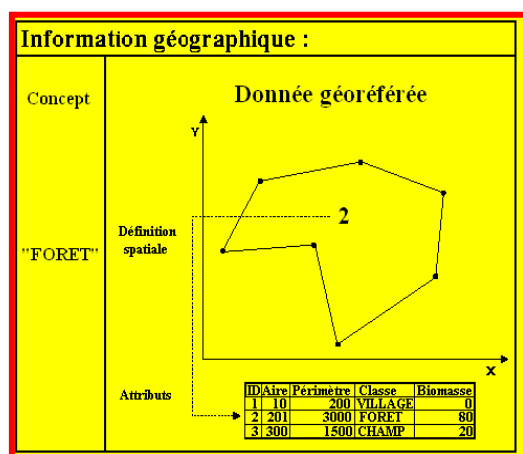


Figure X.1 – Information géographique

X.2 - Le lien entre signifiant et signifié

Il existe deux types de liens : arbitraire ou symbolique. Dans la langue, le lien unissant le signifiant au signifié est arbitraire. Le signe linguistique est arbitraire. Par exemple, le fait que le mot arbre représente la chose "arbre" relève d'une décision arbitraire qui a été entérinée par convention. Par contre, il existe d'autres types de représentations qui montrent un lien

plus logique ou évident avec la chose représentée : le symbole a pour caractère de n'être jamais tout à fait arbitraire, il n'est pas vide, il y a un rudiment de lien naturel entre le signifiant et le signifié. Par rapport au territoire, l'information géographique est un signe parfois arbitraire. Quand par ex. une route est représentée par le mot route. Mais elle peut aussi être de nature analogue à l'objet territorial; ainsi la représentation graphique est véritablement une image forcément réduite du territoire qui reproduit en plus certaines des relations territoriales telles que la distance, la topologie, etc.

On distingue trois niveaux d'abstraction, selon la nature de la formalisation :

1. les modèles iconiques (une photo aérienne) ;
2. Les modèles analogiques (une carte topographique, où les routes sont représentées par des lignes) ;
3. Et les modèles symboliques (modèle mathématique, par exemple).

On a aussi trois niveaux de relation avec l'objet représenté :

1. Les icônes, qui ressemblent à l'objet (ex. un portrait) ;
2. Les indices, en relation directe (en contiguïté), physique avec l'objet (ex : les symptômes d'une maladie) ;
3. Et les symboles, en relation à l'objet en vertu d'une règle prédéfinie, d'une convention ou d'une association d'idée.

XI - Les niveaux de la modélisation géographique

Une deuxième fonction des systèmes d'information géographique est la modélisation du territoire. Modéliser c'est produire une représentation réduite, simplifiée d'un objet. C'est le phénomène de la réduction des connaissances d'un objet spatial. Un modèle peut reproduire les éléments d'un système comme le territoire, mais aussi ses relations; dans sa version la plus aboutie, un modèle débouche sur une simulation de la réalité complètement détachée de celle-ci. Cette fonction de modélisation est particulièrement importante pour les scientifiques, d'autant plus que les modèles de systèmes d'information géographique sont réutilisables et modifiables à l'infini.

XI.1 - Les niveaux de la modélisation géographique

L'analyse de la modélisation appliquée aux systèmes d'information géographique est faite à l'aide de trois types classiques de modèles scientifiques :

1. Les modèles descriptifs ;
2. Les modèles explicatifs ;
3. Et les modèles prédictifs qui s'emboîtent hiérarchiquement dans le sens qu'une prédiction devrait être basée sur une explication elle-même issue d'un modèle descriptif initial.

On sait cependant qu'une prédiction peut être réalisée sans forcément avoir recours à une explication. Par ex. une projection purement statistique de données du passé sur une date future.

XI.2 - Les modèles descriptifs

Ils permettent d'inventorier des éléments du territoire, c'est-à-dire identifier, caractériser et classer ces éléments. Quatre opérations d'abstraction peuvent être appliquées sur les données :

1. Classification : Une abstraction comme la classification est créée quand une ou plusieurs entités sont assignées à une classe d'entité. Une entité est une instance d'une classe d'entité). Une instance est attribuée à une classe si ses caractéristiques observées se conforment à la définition de cette classe.
2. Généralisation : on parle de généralisation quand les définitions de deux classes se confondent, si "vaches" et "moutons" sont tous deux définis comme "animaux", la classe " animaux" est une généralisation de ces deux classes.
3. Agrégation : une agrégation est créée quand des entités de classes identiques ou différentes sont réunies de manière rigide dans une entité plus complexe, comme par ex. des intersections (entités d'une classe) et des tronçons de route (entités d'une autre classe) agrégés forment un réseau routier (entité plus complexe).
4. Et association : ressemble à agrégation mais les entités forment un groupe qui est structuré de manière plus souple. Par ex. les classes "pendulaire" et "voiture" sont associées dans une classe "déplacements pendulaires en voitures".

De tels modèles permettent d'accumuler des données sur un phénomène et de les classer; HUSSY (1996) parle ici de domestication des données. C'est la constitution d'une base de données géoréférencées interrogeable avec des questions du type : quoi ? Et où ? Exemple de modèle descriptif : lorsqu'on réalise une classification d'une image satellite, on classe des pixels selon des concepts d'occupation du sol. Une fois la carte réalisée on peut agréger les classes.

XII.3 - Les modèles explicatifs

Un niveau plus complexe est atteint lorsque les éléments d'un modèle descriptif (abstractions) sont mis en relation et les relations peuvent être de natures très différentes :

- Stochastique : du grec stokhastikos "conjectural", qui est le fruit du hasard, au moins en partie.
- Statistique (boîte noire) : ensemble des techniques d'interprétation mathématique appliquées à des phénomènes pour lesquels une étude exhaustive de tous les facteurs est impossible. Elle aide à la recherche de régularités et peut conduire à la formulation d'explications sur les structures et les processus observés.
- Conceptuel (grey box) : les processus sont connus mais modélisés de manière très simple en incorporant seulement les composantes les plus importantes).
- Déterministe (white box) : "ordre des faits selon lequel les conditions d'existence d'un phénomène sont déterminées, fixées absolument de telle façon que, ces conditions étant posées, le phénomène ne peut pas ne pas se produire". Il s'agit d'une relation causale, c'est-à-dire le fait que la présence d'un élément génère forcément une certaine action ou la présence de certains autres éléments.

C'est évidemment ce dernier cas qui nous intéresse le plus, même si beaucoup de phénomènes ne peuvent être modélisés de cette manière. D'autres relations sont possibles dans un modèle explicatif, mais quand elles n'ont pas la même force en tant que "mécanismes génératifs", elles ne forcent alors pas un concept à avoir une influence sur un autre, et peuvent seulement être considérées comme descriptives. Un modèle explicatif permet par exemple de répondre à une question comme "que se passe-t-il si je modifie tel ou tel paramètre ?".

XI.4 - Les modèles prédictifs

Les modèles explicatifs sont la base pour des modèles prédictifs qui produisent une image d'un état futur. C'est le niveau de la simulation dans le sens où l'on se détache de la réalité pour travailler avec le modèle avec le risque de ne plus rechercher comment la réalité évolue mais comment le modèle va se comporter. Dans la démarche scientifique, la simulation recouvre en fait trois situations assez différentes :

1. Le cas d'une théorie bien établie (donc d'un modèle explicatif éprouvé) et d'une situation actuelle ou future inconnue : la simulation pallie le manque de données empiriques, dans ce cas c'est un outil de prédiction.

2. Dans le cas, où de nombreuses données empiriques sont disponibles mais où la théorie explicative est encore incomplète, la simulation est un outil de théorisation, notamment à l'aide des démarches par essais et erreurs.
3. Un troisième cas est celui où une théorie et des situations réelles sont connues : la simulation prend un aspect didactique, elle est alors un outil de formation (simulateurs de vol, de conduite, etc.).

Seul un modèle déterministe peut prédire à la fois la localisation et la fréquence d'un événement. Ces modèles prédictifs permettent d'établir des scénarios et de les évaluer en fonction de certains objectifs à atteindre pour aide à la décision. Ils aident à répondre à la question "comment atteindre à telle ou telle situation souhaitée ? ».

XII - Un outil d'action pour l'aménagement d'un territoire

Il est évident que les deux fonctions immédiatement visibles que sont la communication et la modélisation impliquent une troisième fonction pratique implicite : en fin de compte, si on communique et on modélise à l'aide des systèmes d'information géographique, c'est bien pour pouvoir agir sur le territoire et l'aménager. Les systèmes d'information géographique sont donc à la fois des outils de connaissance et des outils d'aménagement du territoire.

Cette fonction d'aménagement nous renvoie aux finalités et aux enjeux de la connaissance, qu'elle soit instrumentale ou conceptuelle. L'action de l'homme a principalement favorisé une action à court terme sur le territoire qui s'est surtout traduite par une plus grande efficacité dans l'exploitation des ressources. Au vu des problèmes environnementaux observés aujourd'hui, l'enjeu de la connaissance est la compréhension et la prise en compte des grands équilibres écologiques et socio-économiques nécessaires à la survie des sociétés et de leur milieu.

XII.1 - Les logiques de l'aménagement du territoire

L'aménagement du territoire a trois logiques fondamentales :

1. Logique de limite : maillages territoriaux (extension, diffusion, concentration, souveraineté, etc.) ;
2. Logique de distance : réseaux (centralité, connexité, connectivité, saillance, etc.) ;
3. Et logique d'échelle : hiérarchies (politiques, économiques, culturelles, centralité marginalité, etc.).

Ces trois logiques permettent d'interpréter les éléments territoriaux :

1. les nœuds : le système urbain d'un pays résulte d'une logique d'échelle, de concentration de la population et des activités dans des points particuliers de l'espace pour toutes sortes de raisons économiques, culturelles, politiques.
2. les réseaux : les voies de transport révèlent des phénomènes de centralité, de connexité. La notion même de distance peut être abolie par certains réseaux comme les réseaux informatiques, qui postulent une centralité multiple mais marginalisent dans le même temps ceux qui n'y sont pas connectés. On identifie des réseaux d'énergie, de matière, de biens et services, d'hommes, d'informations et de monnaie.
3. les maillages : la souveraineté d'une entité politique s'exprime par un découpage administratif de l'espace.

Les éléments territoriaux ne sont pas forcément visibles, physiques (découpages administratifs, quoique ceux-ci soient marqués matériellement dans le territoire, les flux d'informations, etc.). On peut même avancer l'hypothèse que les éléments territoriaux les plus signifiants comme les réseaux commerciaux et financiers sont aujourd'hui invisibles à l'échelle du paysage.

XII.2 - Aspects spatiaux : concepts spatiaux

a - Unités spatiales

L'observation et la représentation de la réalité se fait selon trois grands types d'unités spatiales : le point, la ligne, la surface et selon un découpage soit régulier, soit irrégulier comme sur la figure XII.1.

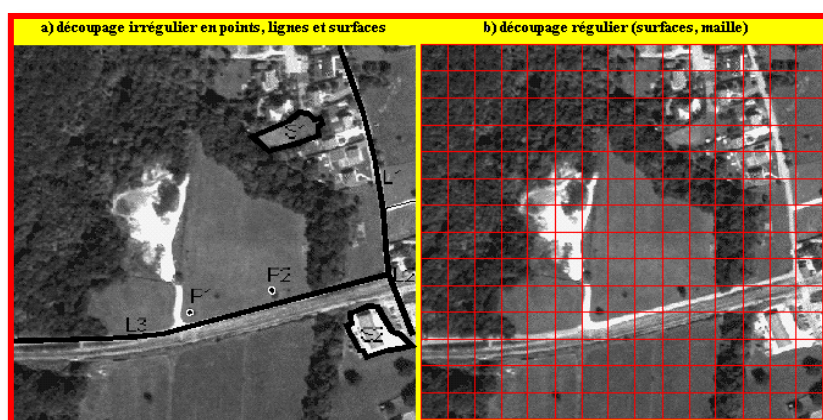


Figure XII.1 – Découpage irrégulier ou découpage régulier

XII.3 - Découpages irréguliers

a - Unités naturelles

Ces unités sont des portions de territoire homogène d'un point de vue naturel, c'est-à-dire biologique, écologique, géologique, etc. On peut ainsi considérer des unités de végétation, des bassins versants, des unités de paysage. Cette approche par unité homogène est centrale dans l'écologie du paysage.

b - Unités administratives

La plupart des statistiques sur l'environnement sont publiées par unité administrative, selon les découpages en vigueur dans chaque pays. Cette approche a deux inconvénients. Premièrement, la hiérarchie du découpage recouvre des espaces très différents suivant les pays et rendent les comparaisons parfois difficiles, ainsi comment comparer le premier niveau sous national des Etats-Unis (les Etats) avec celui de la Suisse (les cantons) ou avec celui de l'Algérie (les Wilayates). Mais surtout, les découpages administratifs ne sont, le plus souvent, pas adaptés aux unités physiques définies par la géologie, la biologie et les autres sciences naturelles.

XII.4 - Découpage régulier : l'approche par maille

Enfin, une dernière approche consiste à définir un découpage a priori, basé uniquement sur des considérations géométriques. C'est le cas type d'un découpage en un quadrillage régulier de cellules (appelées aussi pixels) ayant une taille (résolution) déterminée. Ce découpage a l'avantage de permettre des comparaisons entre des données sur des réalités très différentes (humaines, physiques). Par contre, un tel découpage suppose soit un inventaire exhaustif selon la maille choisie ou soit des méthodes d'interpolation permettant de distribuer les valeurs d'origine selon cette nouvelle grille. Pour la comparaison entre unités spatiales, des problèmes se posent lorsque ces unités ne coïncident pas, ce qui arrive souvent, notamment lorsqu'on utilise des sources de données variées ou des séries chronologiques.

a - Echelle

L'échelle dans les systèmes d'information géographique recouvre deux aspects :

- L'échelle cartographique, soit le rapport entre les distances de la représentation graphique (carte, écran, etc.) et celles de la réalité ;

- Et la résolution, qui définit la taille du plus petit objet observable. Tout objet en dessous de cette taille est ignoré (voir par ex. la résolution d'une image satellitaire).

b - Positionnement

La localisation des objets dans l'espace euclidien, métrique se fait par :

- Référence directe (numérique) : coordonnées géographiques latitude λ , longitude Φ , altitude H ; coordonnées rectangulaires planes X,Y \Rightarrow projection ;
- Et référence indirecte (textuelle) : unité administrative, n° de parcelle, code postal, adresse postale, etc.

c - Relations

Il faut également considérer un espace non métrique, l'espace topologique qui définit les relations spatiales entre les objets. Celles-ci peuvent être implicites (raster, spaghetti) ou explicites (topologie).

XII.5 - Relations entre classes d'objets spatiaux

Le tableau suivant illustre les relations spatiales et topologiques entre objets spatiaux.

	Points	Lignes	Surfaces
Points	Est un voisin Est rattaché	Est proche de Se trouve sur	Est un centroïde de Est à l'intérieur de
Lignes		Croise Joint	Intersecte Est une bordure de
Surfaces			Est couvert par Est adjacent à

XII.6 - Un autre modèle de données : le TIN

Le Triangulated Irregular Network (TIN) est une autre manière de représenter les réalités continues. Ce modèle est à rapprocher du mode raster dans le sens qu'il repose sur une partition de l'espace en un certain nombre de cellules, des triangles de formes irrégulières dans le cas des TINs comme sur la figure XII.2.

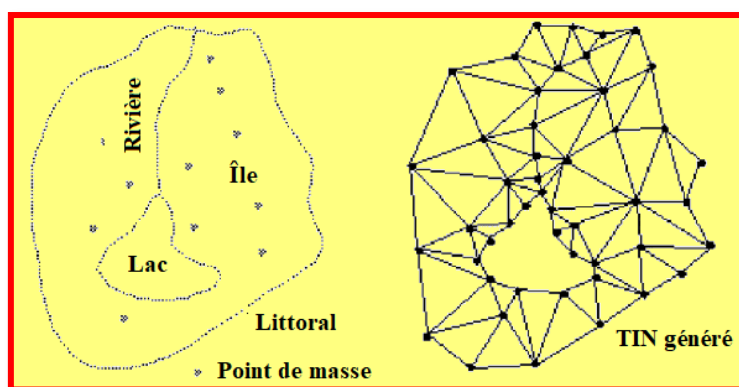


Figure XII.2 – Modèle de des TINs

Le tableau suivant illustre les avantages et les inconvénients des TINs.

Mode	Avantages	Inconvénients
TIN	<ul style="list-style-type: none"> • Sources multiples (points, lignes, polygones) • Réduction du nombre de données • Représente très bien les réalités continues présentant une grande variabilité et de fortes discontinuités (par Ex. terrain) 	<ul style="list-style-type: none"> • Pas idéal pour modélisation de processus spatiaux de surfaces théoriques (coûts, potentiels, etc.). • Difficulté de trouver le nombre et la localisation idéale des points du réseau de triangulation.

XII.7 - Aspects sémantiques : les attributs

a - Echelles de mesure

Les valeurs rattachées aux objets sont situées dans des échelles de mesure suivantes :

- Nominale : classes d'occupation du sol, noms de lieux, etc.
- Ordinale : premier, deuxième ; plus grand, plus petit ; etc.
- Cardinale absolue (origine absolue : température en degrés Kelvin, nb d'habitants, etc.
- Et cardinale relative (origine arbitraire ou conventionnelle : degré Celsius, dates, etc.

b - Attributs thématiques (exemples de la forêt)

b.1 - Indicateurs biophysiques

Un grand nombre d'indicateurs biophysiques permettent de décrire les caractéristiques biologiques et structurales de la végétation, comme :

- L'aire basale : somme des sections des troncs mesurés à hauteur d'homme (1.3 m) sur une surface donnée ;
- La biomasse (du tronc, des branches, des feuilles) ;

- La hauteur (des arbres) ;
- La surface foliaire, exprimée par l'indice foliaire (Leaf Area Index ou LAI) ;
- Le recouvrement : projection verticale des couronnes des arbres sur une surface donnée ;
- La composition floristique : liste des espèces, regroupements en associations ;
- Et la diversité floristique : exprimée par exemple par le nombre d'espèces.

b.2 - Indicateurs socio-économiques

Un certain nombre d'indicateurs permettent de décrire la forêt du point de vue de son utilisation par l'homme :

- Le statut légal : par exemple type de zone d'affectation, type de périmètre protégé (forêt classées, parc, réserve), avec tous les droits et contraintes s'y rattachant ;
- L'utilisation : manière(s) dont est utilisée une forêt (chasse, culture en sous-bois, charbonnage, loisirs, etc.) ;
- Et le niveau d'utilisation, d'anthropisation : niveau d'influence humaine, difficile à mesurer. La FAO distingue par exemple forêt naturelle et plantation.

b.3 - Indicateurs géographiques

Ces indicateurs rendent compte de la distribution spatiale, de la forme ou de la structure d'une unité spatiale :

- La surface : la surface d'une unité, le plus souvent exprimée en hectares dans le domaine forestier ;
- La densité : le rapport entre la surface de forêt et la surface totale d'une unité ;
- Le périmètre : la longueur des bords extérieurs et intérieurs d'une unité ;
- La fragmentation, forme : le rapport entre le périmètre et l'aire ;
- La surface noyau : la surface se situant à une certaine distance du bord de l'unité ;
- Et la proximité : distance à l'unité la plus proche.

c - Aspects temporels

La dimension temporelle souvent difficile à prendre en compte car elle requiert des mesures à intervalles réguliers, les nomenclatures et les méthodes de mesure sont souvent difficiles à comparer dans le temps et la définition spatiale des objets change au cours du temps (par ex. le découpage administratif).

XII.8 – Démarche de construction de la base de données GEOSPATIALE

Le design d'une base de données géographique sera réalisé en suivant la démarche en cinq points suivants (comme sur la figure XII.3) :

1. Définition du modèle conceptuel de la réalité ;
2. Définition des objets et de leur relation (modèle logique) ;
3. Choix des types d'objet spatiaux (point, ligne, polygone) et des modes de représentation (raster, vecteur, TIN) ;
4. Traduction des types d'objets et des modes de représentation dans un logiciel (ArcView ou Mapinfo) ;
5. Et organisation des objets dans la base de données géographique.

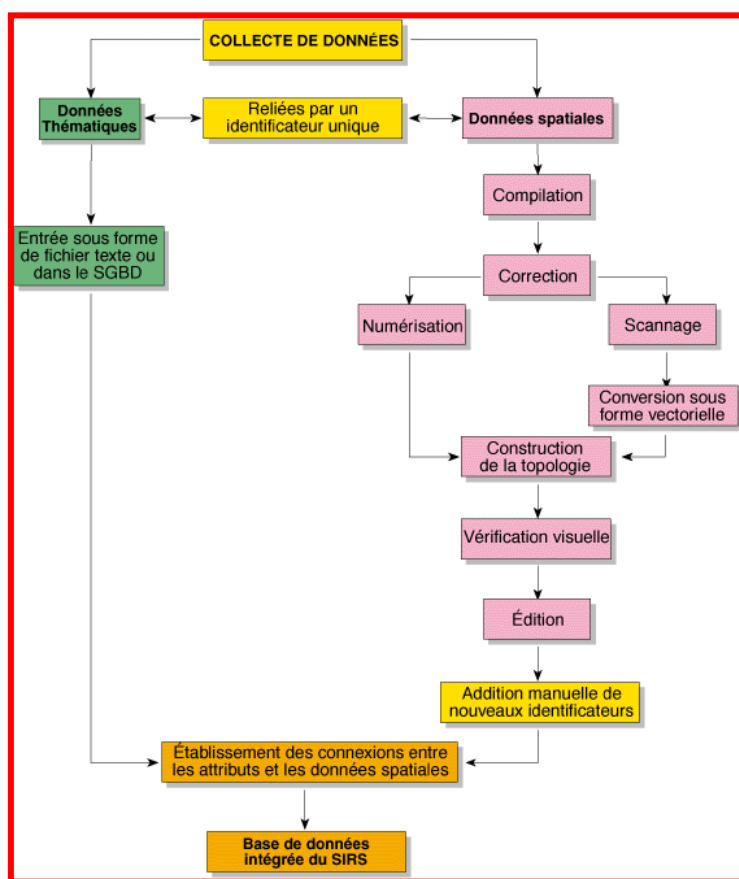


Figure XII.3 – Construction de la base de données GEOSPATIALE

XII.9 - Relations spatiales entre les objets

Les relations spatiales entre classes d'objets (points, lignes, surfaces) distinguent les relations booléennes ("intersecté", "est adjacent", "est inclus", etc.) qui n'autorisent que deux situations (oui ou non), des relations floues (par ex. "est proche de") qui nécessitent la fixation

d'un seuil sur une échelle continue (à partir de quelle distance un objet est-il considéré comme proche d'un autre ?). D'autre part, en mode vecteur, ces relations peuvent être implicites (approche non topologique) ou définies de manière explicite (approche topologique).

XII.10 - Entités spatiales selon l'approche topologique et non topologique

La figure XII.4 illustre la représentation des entités spatiales selon les approches topologique et non topologique.

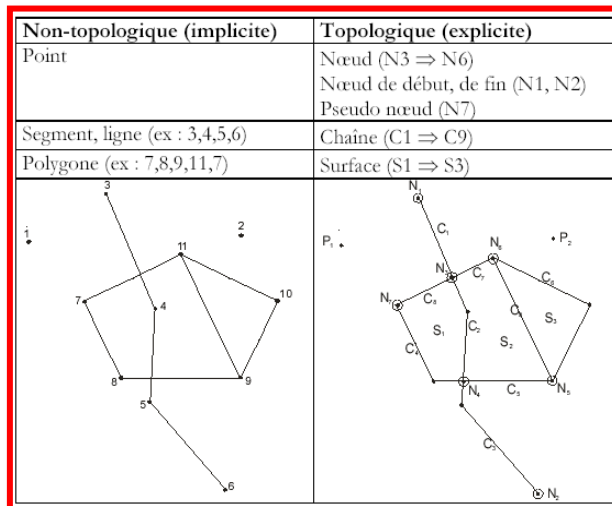


Figure XII.4 – Entités spatiales selon l'approche topologique et non topologique

XIII – Cartomatique et sémiologie graphique

XIII. 1 – Notions de base

La figure XIII.1 illustre les notions de base de représentation des cartes.

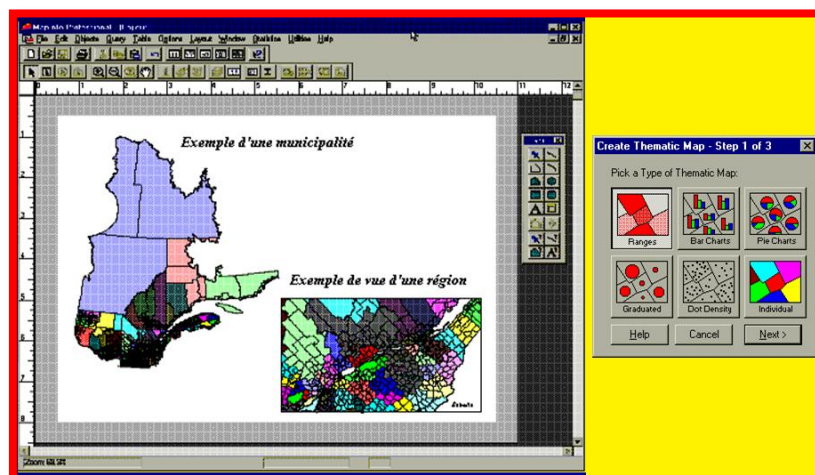


Figure XIII.1 – Notions de base

XIII.2 - Contenu de la présentation

a - Cartomatique et SIRS (Système d'Information à Référence Spatiale)

- Le but des cartes :
 - Inventaire ;
 - Analyse ;
 - Et communication ;
- La sémiologie graphique ;
- Les objets cartographiques et données thématiques ;
- Les métadonnées ;
- Le processus cartographique ;
- Le traitement des données :
 - Thématiques ;
 - Et cartographiques ;
- La représentation cartographique ;
- Et la cartographie dans MapInfo.

La partie informatisée d'un système d'information à référence spatiale comporte deux ensembles complémentaires : les données et les traitements. Les données associent une description géométrique du territoire (géobase) et une description thématique des phénomènes qu'on y retrouve. La nature des traitements est complexe et diversifiée. Pour simplifier, identifions six modules de traitement qui assument des rôles complémentaires : le système de gestion de bases de données thématiques (SGBD), le système de gestion de données localisées (SGDL), les modules de traitement d'images, les modules d'analyse statistique, les modules d'analyse spatiale et les modules de cartographie.

XIII.3 - Objectifs et Qualités d'un SIRS (système d'information à référence spatiale)

On va lister en dix points ci-après les objectifs et les qualités d'un système d'information à référence spatiale :

1. Rentable (en fonction du mandat de l'organisme) ;
2. Pertinent (place l'emphase sur les résultats plutôt que sur les méthodes de production) ;
3. Flexible (capable d'évoluer) ;

4. Sécuritaire (construit de façon modulaire pour minimiser les erreurs et protéger l'information confidentielle) ;
5. Performant (contrôle de la qualité et efficacité des traitements) ;
6. Indépendant (du matériel et des logiciels afin d'en assurer l'évolution et la portabilité) ;
7. Coopératif (bien définir les rôles et les responsabilités des services et des individus impliqués) ;
8. Motivant (impliquer les utilisateurs à toutes les étapes de la conception et de l'utilisation) ;
9. Progressif (éviter les achats massifs d'équipements et limiter la saisie initiale des données afin de réduire les risques d'erreur et les coûts) ;
10. Documenté (rédaction de la documentation parallèlement au développement du système).

Dans ce contexte, on parle de cartographie assistée par ordinateur (CartAO) ou de cartomatique. C'est-à-dire une technique cartographique utilisant un système de traitement informatique comportant un ordinateur et des périphériques spécialisés comme sur la figure XIII.2.

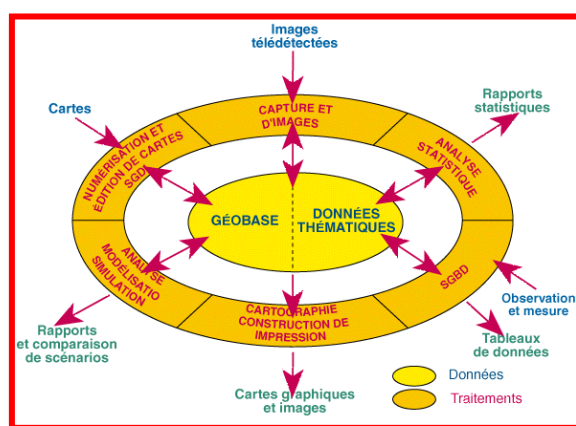


Figure XIII.2 – Vue globale d'un système d'information à référence spatiale (SIRS)

XIII.4 - Quelques domaines d'application des SIRS

La figure XIII.3 illustre les quelques domaines d'application des SIRS (système d'information à référence spatiale) à composante SIT (système d'information territoriale) ou à composante SIG (système d'information géographique).

DOMINANTE SIT		DOMINANTE SIG	
CMU	Cadastré multi-usages <i>Multipurpose cadastre</i>	MNA	Modèle numérique d'altitude <i>Digital elevation model</i>
SGT	Système de gestion des terres <i>Land data systems</i>	SIA	Système d'information pour l'aménagement <i>General information system for planning</i>
SIC	Système d'information cadastrale <i>Cadastral information system</i>	SIF	Système d'information foncière <i>Property information system</i>
SGI	Système de gestion des infrastructures <i>Automated mapping and facility management</i>	SIURS	Système d'information urbaine à référence spatiale <i>Urban information system</i>
		SAIG	Système d'analyse de l'information géographique <i>Geographical information analysis system</i>
		SIEM	Système d'information pour les études de marché <i>Market analysis information system</i>
		SIERS	Système d'information environnementale à référence spatiale <i>Environmental information system</i>
		SIS	Système d'information sur les sols <i>Soil information system</i>
		SGTC	Système de gestion des transports et des communications <i>Transit and communication information system</i>

Figure XIII.3 – Quelques domaine d'application des SIRS

XIII.5 - Inventaire, analyse et communication

La détermination de la fonction de la représentation graphique et l'identification du public auquel elle est destinée sont des éléments essentiels pour prendre des décisions à toutes les étapes de la préparation d'un document cartographique. Selon ces critères on peut distinguer trois types de cartes :

1. Les cartes d'inventaire ou "cartes à lire" montrent avec précision la localisation des phénomènes, qui sont souvent très chargées (plusieurs catégories, plusieurs symboles comme sur la figure XIII.1). Par exemple : cartes pédologiques, cartes d'utilisation du sol, cartes ethniques et linguistiques, etc.
2. Les cartes analytiques ou "cartes à voir" s'intéressent à la répartition spatiale des phénomènes, montrent une vue d'ensemble plutôt que les détails et requièrent souvent un traitement préalable des données.
3. Les cartes thématiques font partie de cette catégorie Les "cartes messages". Ce sont des "cartes à voir" encore plus schématiques. Sont souvent plus proche du schéma que de la carte conventionnelle. Par exemple : les cartes de vulgarisation pédagogique, les cartes publiées dans les journaux.

Ces catégories ne sont pas mutuellement exclusives. Il arrive qu'une même carte puisse se lire à deux niveaux voir même à trois niveaux selon les superpositions et les compositions des couches thématiques des cartes.

XIII.6 - La sémiologie graphique : une grammaire du langage graphique

Pour être compréhensible le langage verbal suit des règles de construction très précises définies par une grammaire. La construction et la compréhension d'une image visuelle repose aussi sur l'application d'une grammaire du langage graphique appelée la sémiologie graphique. Elle regroupe "l'ensemble des règles permettant l'utilisation d'un système graphique de signes pour la transmission d'une information". La construction des symboles cartographiques doit respecter les règles générales de la sémiologie graphique. Le vocabulaire du langage cartographique est constitué par des variables visuelles :

- La position dans le plan ;
- La taille ;
- La valeur ;
- Le grain ;
- La couleur ;
- L'orientation ;
- Et la forme.

Ces variables sont appliquées aux points, aux lignes et aux surfaces, que l'on vient de voir dans les paragraphes précédents. La syntaxe du langage cartographique repose sur le choix de la ou des variables visuelles, possèdent des propriétés perceptuelles aptes à représenter adéquatement les propriétés de l'information géographiques : propriétés de ressemblance, de différence, d'ordre ou de proportionnalité.

XIII.7 - Les objets cartographiques

La cartomatique utilise quasi-exclusivement le mode vectoriel car il supporte la notion d'objets auxquels il est possible de rattacher des attributs thématiques. Les objets sont représentés par des entités géométriques : le point (une borne fontaine), la ligne (un trajet d'autobus) et le polygone (un quartier). Il est souhaitable en cartomatique et essentiel dans un SIRS de tenir compte de la topologie. Celle-ci peut être définie comme une géométrie de situation qui s'intéresse aux relations de voisinage entre les figures géométriques. Ces relations ne sont pas altérées par la déformation des figures donc par les changements

d'échelle ou de projection. Topologiquement, les propriétés topologiques les plus importantes sont l'adjacence, l'inclusion/exclusion, l'intersection et la connectivité. La figure XIII.4 illustre une certaine égalité de forme du point de vue topologique.

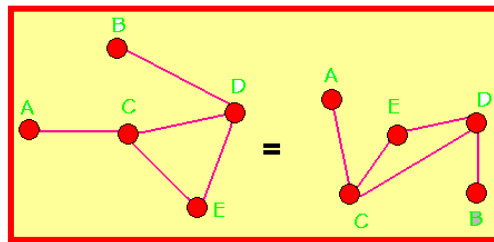


Figure XIII.4 – Topologiquement ces deux formes sont égales

XIII.8 - Une typologie des objets en mode vectoriel

La figure XIII.5 illustre la typologie des objets 0D, 1D et 2D en mode vectoriel.

	Objets définis par leur géométrie	Objets définis par leur géométrie et leur topologie
Objets ponctuels à 0D	Point .	Noeud .
Objets linéaires à 1D	Segment de droite —	Maillon (lien) —
	Ligne brisée	Maillon orienté
	Arc P1	Chaîne (ligne brisée)
	Périmètre (ligne brisée fermée)	Chaîne (arc)
	Périmètre (arc fermé)	Chaîne complète (référence aux noeuds de départ et de fin N1 et N2 et aux zones à gauche et à droite Z1 et Z2)
Objets de surface à 2D	Périmètre formé d'une ligne brisée ou arc ou d'une aire intérieure	Périmètre (séquence de chaînes) et d'une aire intérieure (face)
	Aire intérieure ou face aire (zone ou polygone) sans sa limite	Zone formée de 2 polygones inclus (P1 et P2) et 1 polygone exclus (P3) Z = centroïde de la zone

Figure XIII.5 – Typologie des objets en mode vectoriel

XIII.9 - Les données thématiques

Les données ou attributs thématiques sont des indicateurs qui permettent de qualifier et, plus fréquemment, de quantifier les phénomènes (par exemple le revenu est utilisé pour quantifier le phénomène de pauvreté). Une définition sémantique précise des données est donc

préalable à l'analyse statistique ou à la représentation cartographique des données de même qu'à l'interprétation des résultats découlant de tels traitements. L'échelle de mesure utilisée influence également le type de traitements (statistiques ou cartographiques). L'exhaustivité de la "couverture" (s'agit-il d'un échantillon ou d'une population statistique) est un facteur à considérer.

XIII.10 - Les métadonnées

Il est très important de bien documenter les sources et les traitements effectués sur les données originales (cartographiques ou thématiques). Les métadonnées sont utilisées à cette fin. Elles doivent préciser, entre autres, la référence spatiale, les méthodes de production, les entités et les attributs de même que les organismes impliqués et les droits de diffusion.

XIV – Les phases de création d'une carte thématique

On peut diviser le processus de création d'une carte thématique en plusieurs étapes. Certaines peuvent recourir à des environnements plus spécialisés (les SGDL (système de gestion de données locales) ou les SGBD (système de gestion de base de données) par exemple). Les étapes sont les suivantes (voir la figure XIV.1) :

- Acquisition des informations (Entrée) :
 - Les données originales (cartographiques ou thématiques) proviennent de plus en plus de fournisseurs (souvent des organismes gouvernementaux) ;
 - On doit s'assurer de la compatibilité des formats avec les logiciels utilisés ;
- Traitements :
 - Ils sont fonctions du phénomène à illustrer, du message à transmettre et des caractéristiques des données ;
- Représentation cartographique :
 - Elle doit respecter les règles de la sémiologie graphique ;
- Sortie :
 - Cette étape débouche de plus en plus vers l'infographie.

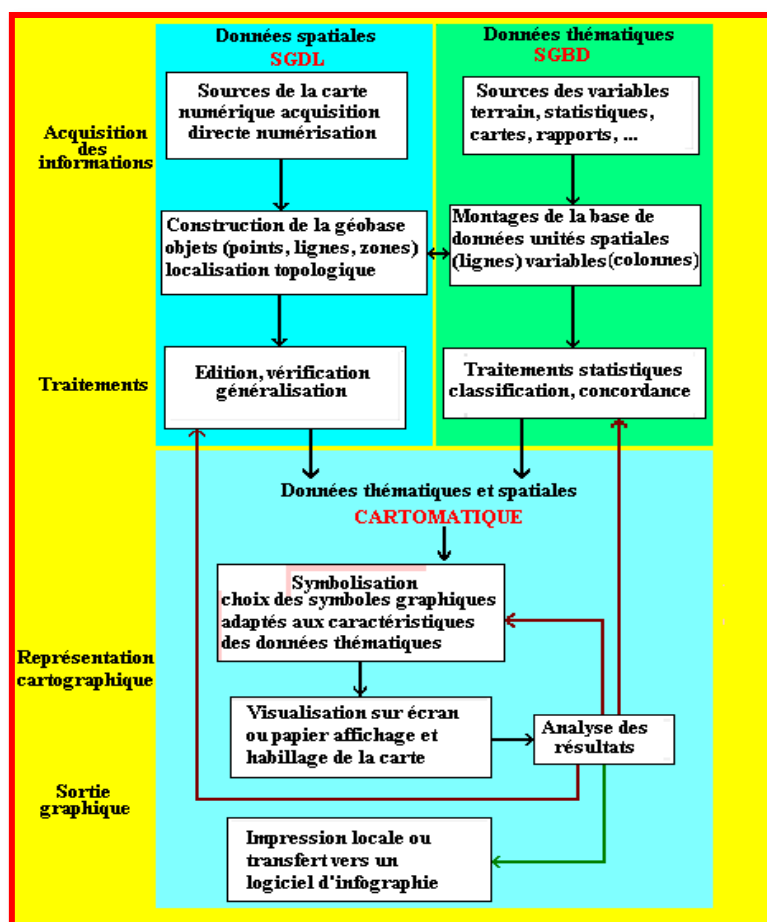


Figure XIV.1 – Création d'une carte thématique

XIV.1 - Traitement des données thématiques

Il est souvent nécessaire d'effectuer un traitement des données thématiques avant la représentation cartographique. La forme la plus élémentaire est l'édition (correction, suppression ou addition des données). On peut agréger les données pour les faire correspondre aux entités spatiales (par exemple, si nos données sont sur la base des municipalités et que notre carte en est une des MRC). Il est aussi possible de créer de nouvelles variables à partir de variables existantes (par exemple la population totale, si elle n'est pas disponible, peut s'obtenir en additionnant les populations féminine et masculine). La classification ou discrétisation est une opération que l'on effectue fréquemment afin de rendre les cartes plus lisibles. Il s'agit de réduire un grand nombre d'observations à un nombre plus petit de groupes en vue d'en faciliter la description. Toute réduction implique une perte d'information. La méthode de discrétisation choisie doit tenir compte de la distribution des fréquences des données originales afin de minimiser cette perte comme sur la figure XIV.2.

Choix d'une méthode de discrétisation selon la forme de la distribution des fréquences

Méthode de discrétisation	Forme de la distribution			
	Unimodale			Plurimodale
	Uniforme	Symétrique (normale)	Dissymétrique	
Amplitude égale en valeur Intervalles égaux	X			X
Fréquences égales Quantités : même nombre d'observations par classe	X			X
Moyennes emboîtées	X	X	X	X
Écart-types		X		
Classes équiprobables écart-types avec fréquence égale par classe	X	X		X
Progression arithmétique, géométrie, réciproque			X	
Empirique (discontinuité) seuls des classes aux discontinuités	X ¹	X ²	X ²	X ³

X : Méthode très appropriée
X : Méthode relative appropriée (selon les caractéristiques de la distribution et les objectifs)

1 . Raveneau (1997) d'après Dumolard, 1982 et Evranx, 1977; modifié
2 . S'applique bien aux distributions comportant un petit nombre d'observations (moins de 50)
3 . L'observation des discontinuités peut aider également au choix du nombre de classes

Figure XIV.2 – Choix des méthodes de discrétisation

XIV.2 - Traitement de la géobase

Il est possible d'effectuer plusieurs traitements sur la géobase (base de données géographiques) par exemple :

- Édition simple
 - Élimination ou ajout d'objets cartographiques
 - Modification des objets existants
- Agrégation
 - Création de nouveaux objets par regroupement d'objets existants
- Généralisation
 - Simplification des objets
 - Souvent nécessaire si la géobase provient d'un fournisseur et qu'elle a été conçue pour un usage à grande échelle qui demande plus de précision

La figure XIV.3 donne un exemple du résultat de cette opération.

	Elimination d'objets	Lissage	Abstraction de la forme	Regroupement d'objets
Carte à grande échelle				
Carte à petite échelle				

Figure XIV.3 – Représentation des objets à grande et à petite échelle

XIV.3 - Discrétisation et représentation cartographique

L'exemple de la figure XIV.4 démontre comment l'utilisation de différentes méthodes de discrétisation peut influencer le résultat final et, par conséquent, le message que transmet la carte. Dans la première carte (a), le chercheur a fixé les limites de classes en examinant la distribution des valeurs. Dans la seconde carte (b), il utilise une méthode qui divise l'étendue en intervalles égaux et enfin Dans la troisième (f), il retrouve le même nombre d'entités (municipalités) dans chacune des classes. La question que l'on se pose est : quelle est la meilleure méthode ?

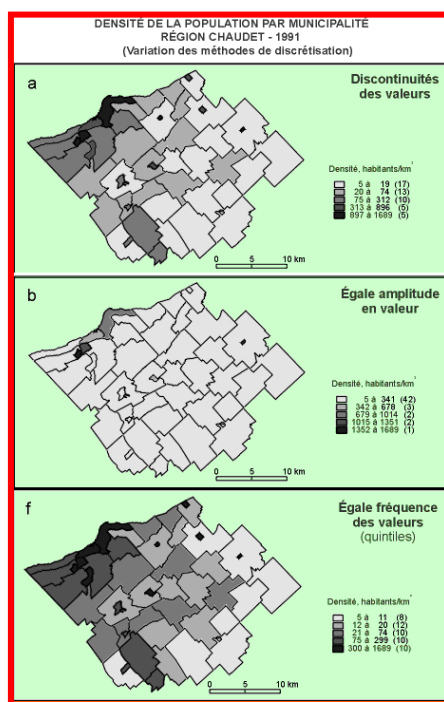


Figure XIV.4 – Variation des méthodes de discrétisation

XIV.4 - Le processus cartographique dans MapInfo

Le processus s'effectue en plusieurs étapes : le choix de la méthode de représentation en fonction des caractéristiques des données, le choix de la variable à cartographier, sachant que la variable peut être le résultat d'un calcul, et la variable peut provenir d'une ou plusieurs tables. L'ajustement des différents paramètres, qui est liée à la méthode retenue (ici une carte choroplète) comme sur la figure XIV.5.

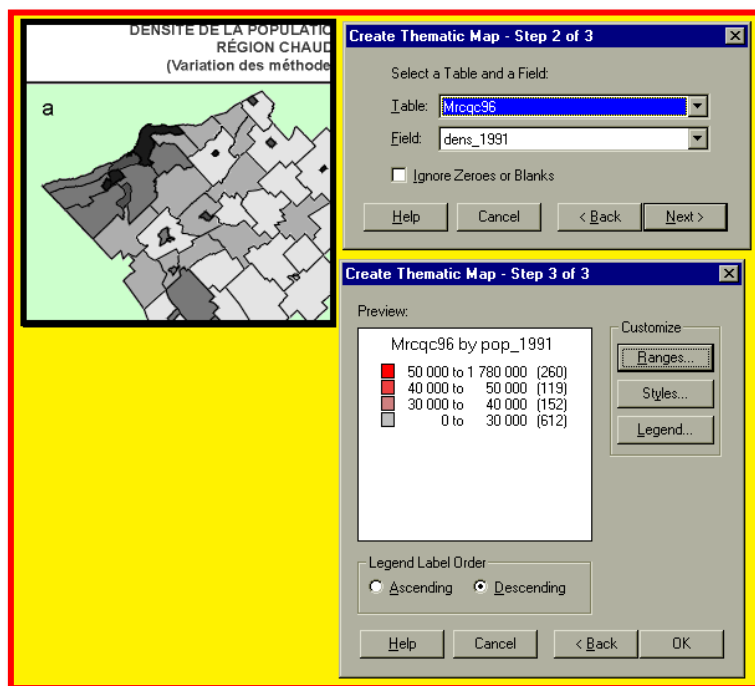


Figure XIV.5 – Différentes représentations sous MapInfo

XIV.5 - Un exemple dans MapInfo La population d'un centre urbain en 1991

Plusieurs types de cartes sont réalisables dans MapInfo. On voit ici comment la distribution de la population peut être représentée en utilisant des symboles proportionnels, un nuage de points (dot density) ou alors une carte choroplètes (trames). Notez que ces exemples sont incomplets (pas de titre, d'échelle ou de source) et ne sauraient être considérés comme des produits finis comme sur la figure XIV.6.

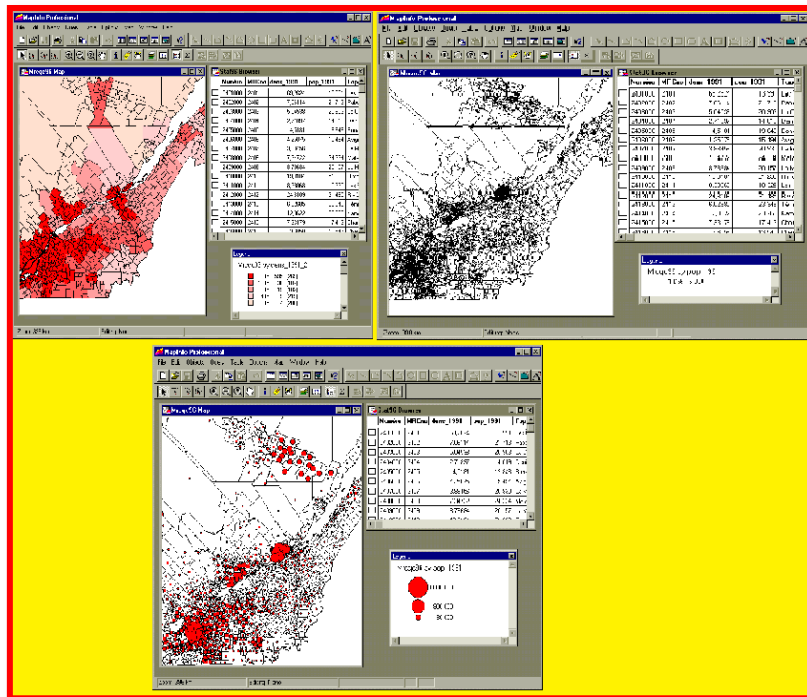


Figure XIV.6 – Différentes cartes urbaines sous MapInfo

XV – Généralités sur les bases de données géographiques

Il existe un grand nombre de système d'information géographique sur le marché. Tous n'ont pas la même architecture ni les mêmes fonctionnalités. Peut-on définir un système d'information géographique complet ? Les composantes d'un système d'information géographique sont :

- Une base de données à caractère spatial et thématique ;
- Un système de représentation cartographique ;
- Un système de saisie numérique ;
- Un système de gestion des bases de données géographiques ;
- Un système d'analyse spatiale ;
- Un système de traitement d'images ;
- Et un système d'analyse statistique.

a - La base de données à caractère spatial et thématique

Elle constitue le noyau du système et se compose d'objets spatiaux (fonds de carte) et de leurs attributs. Elle diffère selon qu'elle concerne les deux modes vecteur ou raster :

- Mode raster : Les deux parties de l'information géographique forme+localisation et attributs constituent une seule composante.
- Mode vectoriel : La base de données à caractère spatial et les attributs sont séparés et constituent 2 composantes.

On note que la base de données attributaires est souvent gérée par un logiciel autonome.

b - Le système de représentation cartographique

Ce système est le plus élémentaire et tous les logiciels le proposent. Il dépend des bases de données spatiales et attributaires (sélections, conditions, etc.). On remarque que les logiciels de système d'information géographique ont trop souvent des possibilités minimalistes en matière de cartographie. On voit que par exemple que MapInfo est plutôt bien loti de ce point de vue ainsi que ArcInfo ou ArcView.

c - Le système de saisie numérique

Cela comporte les possibilités de saisie de cartes selon des formats plus ou moins convertibles, et les capacités d'importation du logiciel d'images, de données textuelles formatées, etc. Par exemple, l'importation de tables au format texte ou Excel par MapInfo.

d - Le système de gestion des bases de données géographiques (SGBDG)

Le SGBD est chargé de la gestion des attributs thématiques des objets spatiaux. La particularité d'un SGBD dans un SIG est l'existence d'outils gérant les attributs et les objets spatiaux conjointement. Cela permet d'introduction de nouvelles informations thématiques, d'interrogation sous différentes conditions, d'analyse thématique et/ou spatiale et d'assurer la production de résultats (textes ou cartographiques) cohérents.

e - Le système d'analyse spatiale

Ce système donne une des particularités fondamentales des systèmes d'information géographique, qui permette une analyse basée sur les caractéristiques spatiales des données. Par exemple, la mise en évidence de zones propices à une exploitation forestière, avec des contraintes hydrographiques, pédologiques (qualité du drainage), végétales (présence d'essences plus ou moins intéressantes), foncières (coût des terrains concernés par la zone d'étude) : Les requêtes sont fondées sur la superposition des objets spatiaux et leur mise en relation ensembliste (selon les opérateurs de l'analyse combinatoire : intersection, union,

complément de l'intersection, complément de l'union, union exclusive). Ce système utilise les notions de base de la géographie telles que la localisation, la distance, la connexité et la superposition. Il se nourrit des bases de données spatiales et thématiques et les enrichit. On remarque qu'en mode vectoriel, ce système devient problématique sans les informations topologiques.

f - Le système de traitement d'images

Ce système permet l'analyse d'images numériques de télédétection, il doit offrir un large éventail de méthodes globales et locales.

g - Le système d'analyse statistique

Ce système doit autoriser des analyses plus ou moins complexes sur les attributs mais aussi sur les objets spatiaux.

h - Les composantes d'un SIG : Apports et contraintes de l'information géographique numérique en mode raster et en mode vecteur

h.1 - Le mode raster

On va voir les avantages et les inconvénients du mode raster.

- Avantages : Le mode raster présente les objets spatiaux de façon continue, à l'image de la réalité du terrain. Cela le rend particulièrement adapté à l'étude de nombreux phénomènes naturels : altitude, précipitations, biomasse, etc. Par sa structure, il se combine facilement aux images satellitaires et aux photographies aériennes.
- Inconvénients : La taille des fichiers numériques pose des problèmes d'intégration et de délais de traitements, les images sont assez éloignées des critères de représentation cartographique, et le transfert des données contenues dans une image vers la carte demande un travail spécifique difficilement automatisable (voir la figure XV.1).

Pour le mode raster, la valeur des pixels, en mode maillé, c'est la valeur du pixel qui traduit la relation entre le mode sémantique et le mode géométrique. Chaque pixel possède au minimum une valeur de luminance, qui traduit la quantité de lumière qu'il reçoit du sol. On peut associer un vecteur de grande dimension si l'on dispose des informations nécessaires (luminance, altitude, coordonnées, localisation relative, nature du sol, etc.). Les valeurs des

pixels sont discrètes (exemple : dans le cas d'une valeur de luminance non traitée) ou classifiées (exemple : dans le cas d'un résultat de traitement statistique par ACP ou AFM).

h.2 - Le mode vectoriel

On citera bien sur les avantages et les inconvénients du mode vectoriel. Il est très performant sur les espaces discrétisés, puisqu'il décrit lui-même (de part le modèle du graphe planaire) l'espace en composantes élémentaires de dimensions zéro, un ou deux : plans d'urbanismes, cadastre, etc. Il est aussi très performant pour l'étude des réseaux (idéalement décrits par des vecteurs, linéaires et orientés). La relation entre les objets spatiaux et leurs attributs est nette et les requêtes sont facilitées. Sa structure est proche de celle de la cartographie traditionnelle (voir la figure XV.1). En ce qui les inconvénients, c'est qu'il est adapté aux seuls espaces discrets (par opposition aux phénomènes de champs).

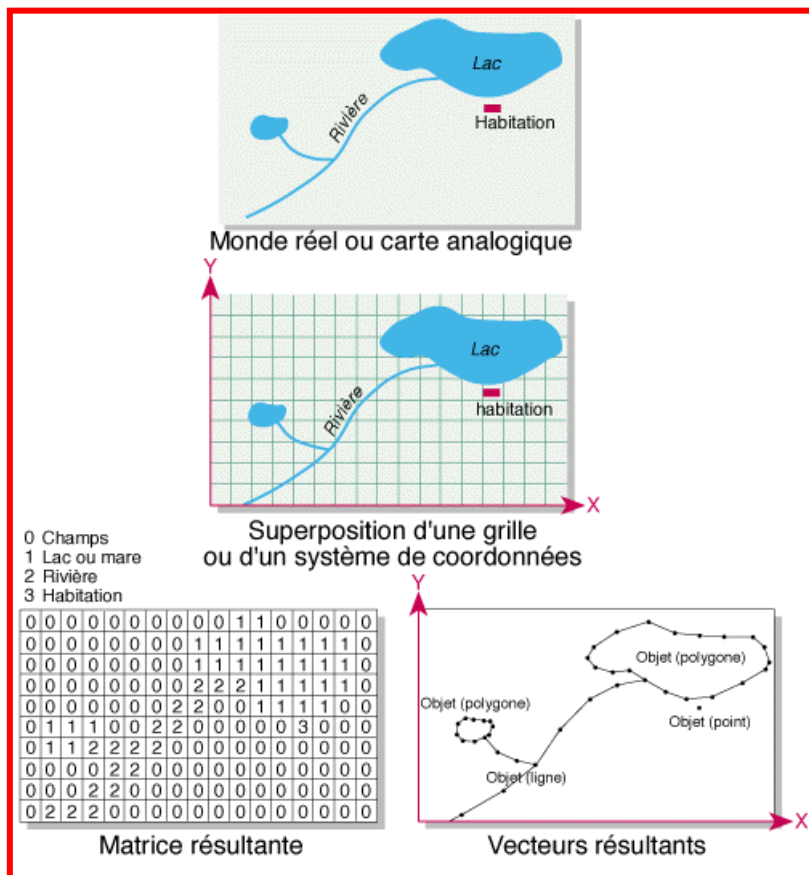


Figure XV.1 – Raster ou vectoriel

Pour le mode vecteur, Les relations de construction et de composition permettent la représentation des entités du monde réel se fait à partir de trois figures de base : le point, le segment de droite, le polygone Ainsi, on construit une rivière à partir d'arcs, de sommets et de

faces : on parle de relation de construction entre la rivière (objet « sémantique ») et les figures géométriques utilisées. Certains objets complexes sont composés d'objets plus simples, eux-mêmes construits comme précédemment. On parle alors de relation de composition (exemple d'une ville) comme sur la figure XV.2.

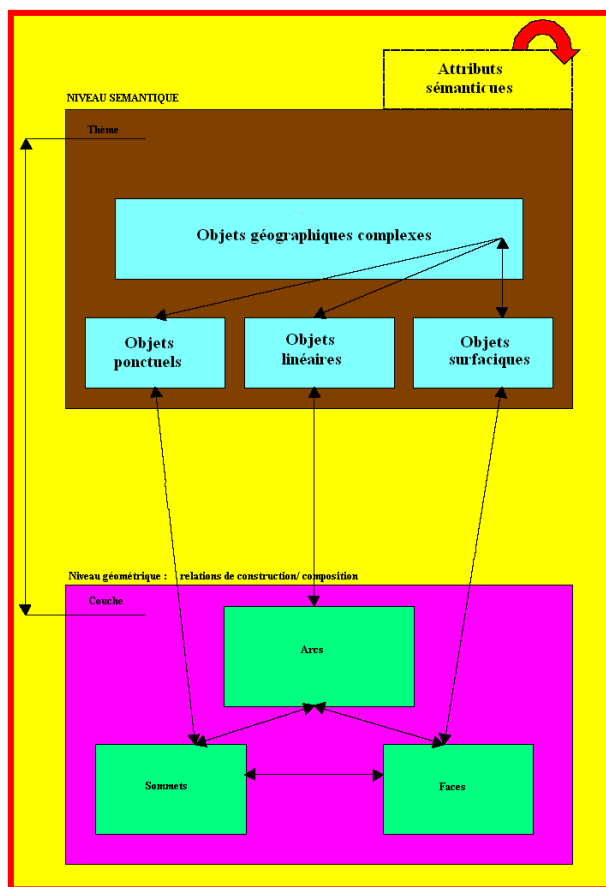


Figure XV.2 - Modèle conceptuel de l'information géographique

XV.1 - Architecture des bases de données géographiques

a – Définitions

Une base de données géographique sur une région d'étude quelconque est organisée sous forme de cartes empilées. Chaque élément de la base est appelé couverture ou coverage en mode vecteur, et couche ou layer en mode raster.

b - Les couvertures

Elles représentent un seul type d'objet spatial (parcelles, îlots, etc.) qui peut être associé à une série d'attributs.

c - Les couches

Elles représentent la variation spatiale continue d'un seul phénomène (champs) ou d'une seule variable (ex : altitude exprimée dans un modèle numérique de terrain). Il est possible d'associer les objets spatiaux des couches à une table d'attributs.

d - Géoréférencer

Il s'agit d'une méthode qui consiste à définir toutes les couvertures et les couches dans un système de localisation (par coordonnées) unique. Le géo-référencement est un préliminaire indispensable à toute forme d'analyse spatiale.

XV.2 - La base de données géographique : l'analyse

La constitution d'une base de données géographique en couches ou en couvertures est avant tout nécessaire à l'analyse spatiale, qui fonctionne beaucoup par la mise en évidence de corrélations. Elle en constitue aussi les limites.

a - Outils d'analyse proposés

Selon les capacités des différents logiciels, on peut distinguer quatre catégories :

1. Requêtes ou interrogations ;
2. Algèbre de carte ;
3. Opérateurs de distance ;
4. Et opérateurs contextuels

a.1 - Interrogation de la BDG

Les requêtes peuvent être simples : récupération d'une information existante (coordonnées géographiques d'une ville, altitude, etc.) ou plus complexes, avec des conditions. On peut avoir des questions suivantes comme exemples :

1. Quelles zones ont la meilleure aptitude agricole, pour telle catégorie d'exploitation ?
2. Quelles sont les zones marécageuses de plus d'un Ha et qui sont adjacentes à des zones industrielles ?

Ces exemples supposent que le système d'information géographique puisse appliquer des opérateurs logiques aux objets géométriques concernés, ici des polygones (outils de l'analyse combinatoire). On remarque que les logiciels, qui procèdent par étapes, on doit

constituer des couches ou des couvertures intermédiaires avant d'exprimer le résultat final (exemple de l'analyse spatiale sur les zones propices à l'exploitation forestière ou autres choses).

a.2 - Outil d'analyse thématique : l'algèbre de carte

Ces outils sont fondés sur la combinaison des couches avec des opérateurs mathématiques. Par exemple, l'utilisation d'un modèle de régression pour l'estimation de la température en fonction de l'altitude et l'estimation du potentiel d'érosion fondée sur la pente et l'intensité de la pluviométrie. On peut distinguer trois types d'analyses :

1. L'ensemble des valeurs de la couche est modifié par un opérateur avec une constante (opération scalaire) ;
2. L'ensemble des valeurs de la couche est modifié par une opération standard (fonction log, etc.) ;
3. Et la combinaison mathématique de différentes couches par +, -, *, /, pour produire des couches – résultats.

On remarque que la troisième opération est basée sur la superposition des couches mais elle utilise une combinaison mathématique et non logique, comme dans les requêtes.

a.3 - Les opérateurs de distance

Ces outils permettent d'intégrer la notion de distance dans l'analyse. On distingue :

- Les zones tampons (buffer) ;
- L'accessibilité ;
- Et les effets de barrière ou frictions

Le système des zones tampons (buffer) permet de créer un couloir autour des objets spatiaux. Par exemple, le couloir de nuisance d'une autoroute ou d'une zone industrielle. Te le couloir non constructible autour d'une zone à risques, etc. Le système d'accessibilité est analysé par les mesures des plus courtes distances ou d'autres modèles plus complexes.

Dans les effets de barrière ou frictions, on analyse la distance par coût, effort, temps (exemple des cartes isochrones) selon le coût que la distance introduit, on détermine des zones à coûts différenciés : il s'agit de discrétiser un effet de champs. Par exemple, le tracé d'une route, les facteurs de localisation d'une entreprise, etc.

a.4 - Les opérateurs contextuels (fichiers raster)

On les appelle aussi opérateurs de voisinage ou opérateurs locaux, car le résultat est obtenu par des transformations des valeurs des pixels d'une couche en fonction de leur environnement (les valeurs des pixels voisins : on parle aussi d'analyse matricielle).

b - Systèmes complexes

Il s'agit de mettre en évidence des phénomènes dynamiques, intégrant éventuellement le facteur temps, et non simplement des situations statiques. Ils nécessitent la constitution de bases de données de très grandes tailles et l'intégration d'outils d'analyses plus performants.

c - Modélisation

Cette démarche se libère de la réalité brute au profit de la virtualité, au double sens potentialité et simulation du mot. Les processus étudiés sont considérés en termes de scénario, et surtout de répondre à la question et si ? On peut citer un simple exemple, les conséquences de l'élévation du niveau de la mer sur le tracé d'une côte.

Un système d'information géographique ne peut être établi que par une équipe pluridisciplinaire. La place du géographe, capable de développer les processus spatiaux et d'optimiser le système d'information géographique doit être essentielle.

XVI – Les systèmes de projections et les référentiels

La géodésie est la science de la forme et la dimension de la Terre et de son champ de pesanteur. Pour représenter et localiser une entité géographique, il nous faut savoir la forme de la Terre et les moyens de se repérer à sa surface. Alors, il faut des systèmes de références comme sur la figure XVI.1.

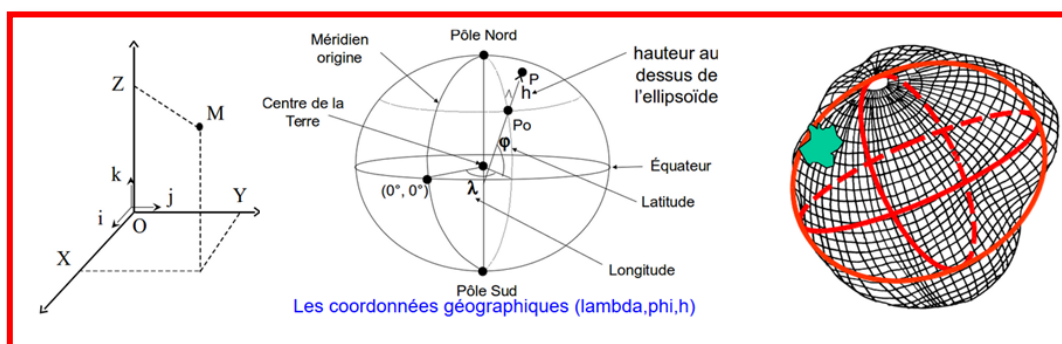


Figure XVI.1 – Les systèmes de projections et les référentiels

Il existe plusieurs systèmes de projections comme sur la figure XVI.2. on a la représentation cylindrique directe, la représentation cylindrique transverse, la représentation conique directe et la représentation conique oblique.

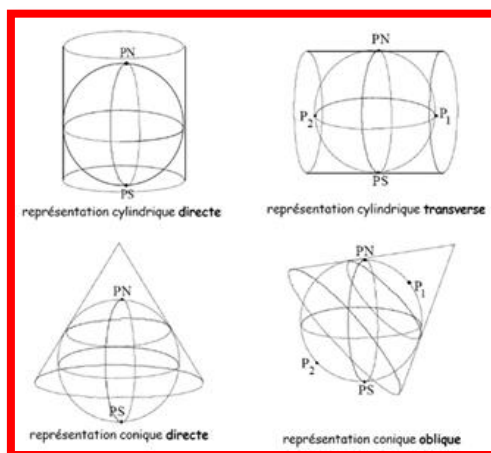


Figure XVI.2 – Les différentes représentations de projection

Toute projection induit des distorsions : le long de la ligne tangente (distorsion nulle) et plus on s'en éloigne, plus la distorsion augmente comme sur la figure XVI.3.

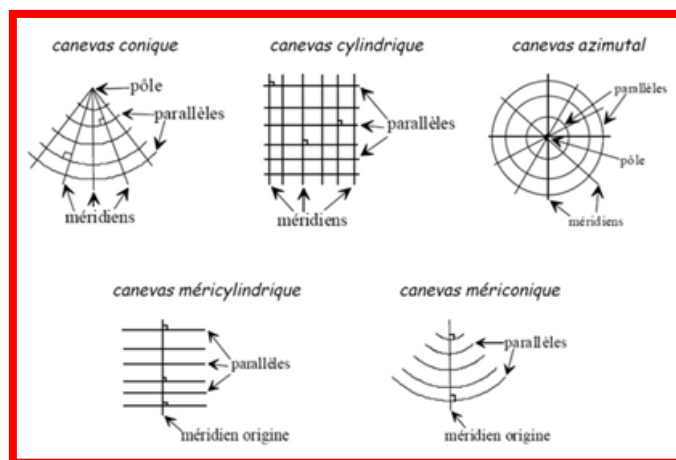


Figure XVI.3 – Les distorsions des projections

Une projection peut conserver les angles (projection conforme), les surfaces (projection équivalente) et ni l'un ni l'autre (projection aphyllactique). Mais, on n'aura jamais les deux à la fois. Les systèmes de projections sont consignés sur la XVI.4. On peut citer les deux normes connues suivantes :

1. World Geodetic System 1984 (WGS 84) est un système associé au GPS basé sur le principe de l'ellipsoïde.

2. Le réseau géodésique français 1993 (RGF 93) est un système basé sur le principe de la densification de l'ETRS 1989. Les présentations Lambert I, II, III, IV et II étendu correspondant au système géodésique NFT. Pour disposer d'une projection correspondant au système de référence RGF 93, on définit la projection dite Lambert 93 de type conique.

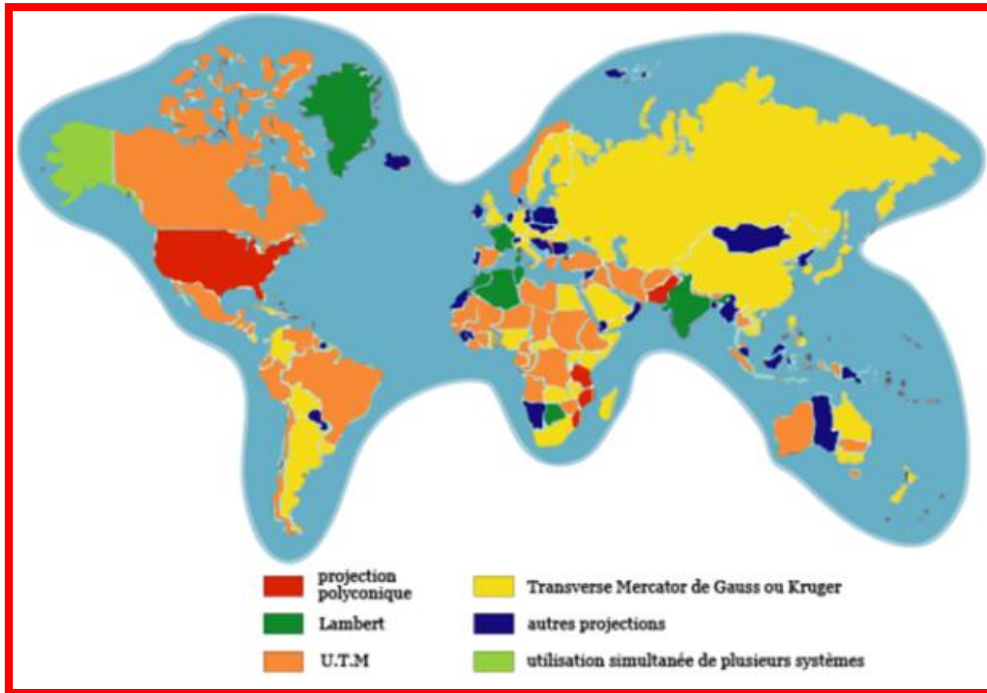


Figure XVI.4 – Les systèmes de projection dans le monde