Chapitre : 02 Représentation des données dans les SIG

1. Introduction

Lorsqu'on évoque les logiciels de cartographie informatique ou les SIG, la première chose qui vient à l'esprit est la gestion des données. Cela suscite diverses questions : Comment est-il possible d'enregistrer les données présentes sur une carte sur un ordinateur? Comment les SIG traitent-ils les données? Dans ce chapitre, nous aborderons ces questions et nous focaliserons sur les différentes techniques d'acquisition des données dans un SIG.

2. Quelques définitions

- **Entité**: Une entité est un phénomène d'intérêt que l'on retrouve dans la réalité le monde réel) et que l'on ne peut plus par la suite le subdiviser en phénomène de même genre.

Exemple: A l'inverse, un massif forestier (entité) peut être subdiviser en massif forestier de plus petite taille. Un type d'entité correspond à tout groupe de phénomènes similaires à être représentés de manières uniforme.

Par exemple les routes (ligne), les rivières (ligne), les altitudes (point), la végétation (zone). Pour chaque type d'entité, il faut produire des définitions précises.

- **Objet:** Un objet est la représentation numérique (ou géométrique) de l'entité. La méthode de représentation numérique d'un phénomène varie en fonction de l'échelle, des besoins et de plusieurs facteurs.
- **Attribut** : Les attributs sont les caractéristiques de l'entité.

Exemple: Entité: route, les attributs de la route: nom de la route (Chemin de wilaya CW, Route nationale RN, largeur de la route, longueur de la route, état de la route, …ect).

- **Notion de couche de donnée** : Une couche est un plan réunissant des éléments géographiques de même type. On dit qu'il s'agit d'un plan car les données géographiques sont bidimensionnelles (x,y).

3. Modes d'acquisition de données

En général, les données géographiques proviennent de différentes sources et sont obtenues à l'aide de différentes méthodes, ce qui les rend multi sources. Il est possible de les recueillir directement sur le terrain à l'aide de relevés topographiques, de les capturer à distance à l'aide de technologies comme le GPS, les photographies aériennes ou les images satellitaires, ou encore d'extraire des cartes et plans existants. La production de données géographiques est également assurée par des organismes spécialisés qui fournissent des informations qui sont ensuite intégrées. Il est nécessaire d'harmoniser ces données, même si elles diffèrent en termes

de qualité, de fiabilité, de précision et de couverture spatiale. Dans la suite de ce texte, nous examinerons les principales techniques d'obtention de ces données.

3.1 Numérisation

La numérisation, également connue sous le nom de digitalisation ou de vectorisation, consiste à prendre en compte la géométrie des objets présents sur un plan ou une carte déjà existant. Cette méthode consiste à déplacer un curseur sur un plan fixé à une table à numériser, préalablement calibrée en coordonnées. Les signaux électriques émis par le curseur sont sensibles à la table, qui peut les localiser sur son plan avec une précision pouvant aller jusqu'au dixième de millimètre.

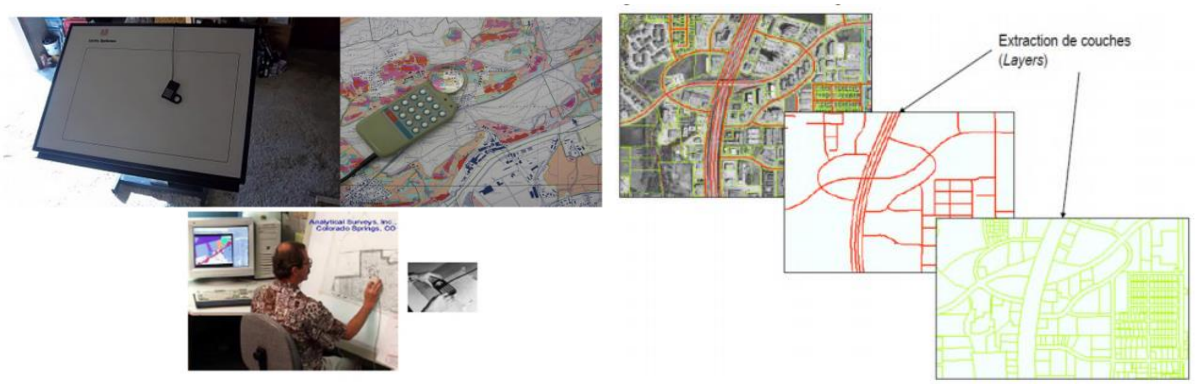

Figure II.1: Les étapes de la numérisation

3.2 Balayage électronique

Un autre moyen de saisir un plan existant est le balayage électronique (effectué à l'aide d'un scanner), qui est plus rapide que la digitalisation manuelle.

Figure II.2: Carte numérisée

3.3 La photogrammétrie

On utilise fréquemment la photogrammétrie aérienne pour réaliser des cartes à échelle moyenne et elle est également privilégiée dans les pays où la couverture cartographique et géodésique est limitée. Elle offre également la possibilité de concevoir des plans à grande échelle à un prix qui peut s'avérer très abordable.

Figure II.3: Exemple des photos aériennes

3.4 La télédétection

La télédétection représente une approche extrêmement pratique pour produire des informations à incorporer dans les Systèmes d'Information Géographique (SIG). Elle implique l'utilisation de photographies aériennes ou d'images capturées et transmises par satellite dans des conditions spécifiques et rigoureuses.

Figure II.4: Exemple d'une image satellitaire

4. Données et information

Généralement pour qu'un objet spatial soit bien décrit et prêt à être utilisé par un SIG, trois informations doivent être fournies:

- Sa position géographique dans l'espace
- Sa relation spatiale avec les autres objets spatiaux : topologie
- Son attribut, c'est à dire ce qu'est l'objet avec un caractère d'identification (code).

Les systèmes d'information géographique permettent de traiter les données spatiales et associées

4.1 Données dans un SIG

Les données dans un SIG sont des données géospatiales ou géographiques, c'est-à-dire des informations liées à des positions ou des caractéristiques spatiales. Les deux principaux types de données dans un SIG sont :

-des données permettant spatial (ou géométriques) de connaître la forme et la localisation dans l'espace de ou des objets observés, En langage SIG, on appelle ces données des « couches géographiques » (Vecteur, raster)

-des données alphanumériques permettant de décrire les objets géographiques, d'en connaître des caractéristiques qualitatives et/ou quantitatives, En langage SIG on les appelle des données « attributaires »

4.2 Information dans un SIG

L'information dans un SIG est obtenue lorsque les données géographiques sont traitées et analysées pour répondre à des questions ou fournir des solutions à des problèmes spécifiques. Ce processus transforme les données brutes en informations exploitables, permettant aux utilisateurs de comprendre des phénomènes spatiaux et de prendre des décisions informées.

5. Types de données dans un SIG

Dans un SIG, il est primordial de comprendre les concepts de données liées, de classification, d'identification et d'attributs afin d'organiser, d'analyser et de comprendre les informations géospatiales.

5.1 Données associées

La représentation géométrique de l'entité spatiale est enrichie par les données liées aux entités géographiques. Il y a un code d'identification pour chaque élément spatial, numérique ou alphabétique, qui sert d'étiquette descriptive pour le point, la ligne ou le polygone en question. Il existe différentes catégories de données qui peuvent être classées :

- Données de classification ;

- Données d'identification ;
- Données attributaires.

5.1.1 Données de classification

Ces informations permettent de classer les points isolés, les lignes ouvertes ou fermées, dans des catégories particulières comme les limites administratives, les contours de parcelles, les bordures de trottoirs, les alignements d'arbres, les conduites de réseaux d'eau, etc. Ces différences sont fréquemment incorporées dans la structure du processus de saisie. En quelque sorte, il s'agit de différencier plusieurs couches d'informations, qui sont prises successivement.

5.1.2 Données d'identification

Les informations d'identification dans un SIG jouent un rôle crucial en permettant de différencier et de caractériser de façon distincte chaque entité représentée sur le plan. Ces données comprennent des informations particulières comme le nom propre de l'objet, comme le nom d'une commune, ou des identifiants numériques permettant une reconnaissance précise, comme le numéro d'une parcelle cadastrale ou celui d'une vanne. De cette manière, ces données rendent la gestion, l'analyse et le suivi des objets géographiques plus simples en les rendant accessibles et manipulables dans le cadre de différentes applications SIG.

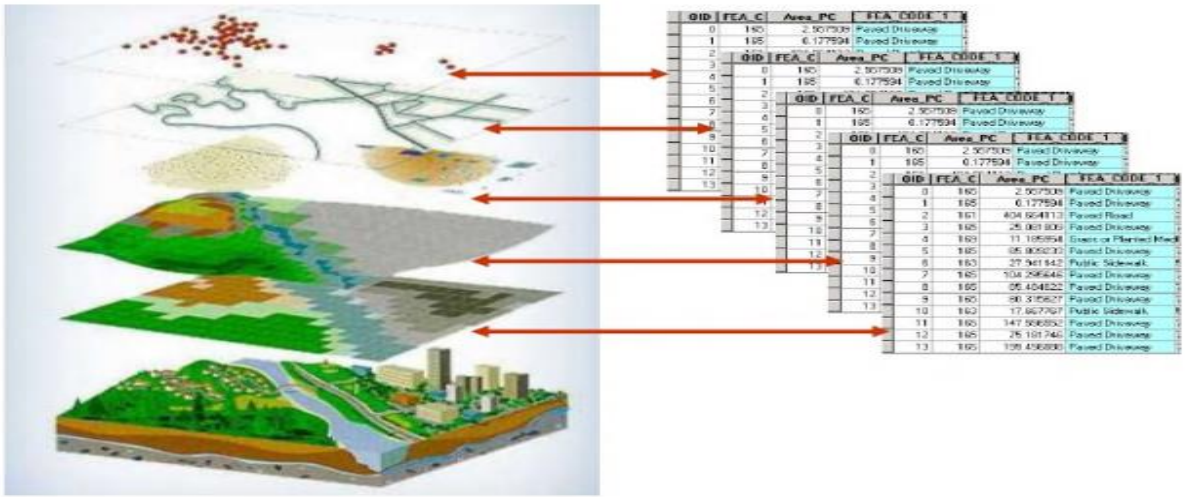

Figure II.5: Exemple des données d'identification dans un SIG

5.1.3 Données attributaires

Les informations fournies par ces données sont spécifiques à chaque objet identifié, telles que le propriétaire de la parcelle, le diamètre de la conduite d'eau, le nom et les coordonnées des

stations pluviométriques. Ces données sont souvent déjà présentes dans des fichiers informatiques, où elles sont associées à l'identifiant de chaque objet. La classe d'objet est généralement déterminée, au moins en partie, par le processus de numérisation, les identifiants étant souvent ajoutés en bloc à la fin.

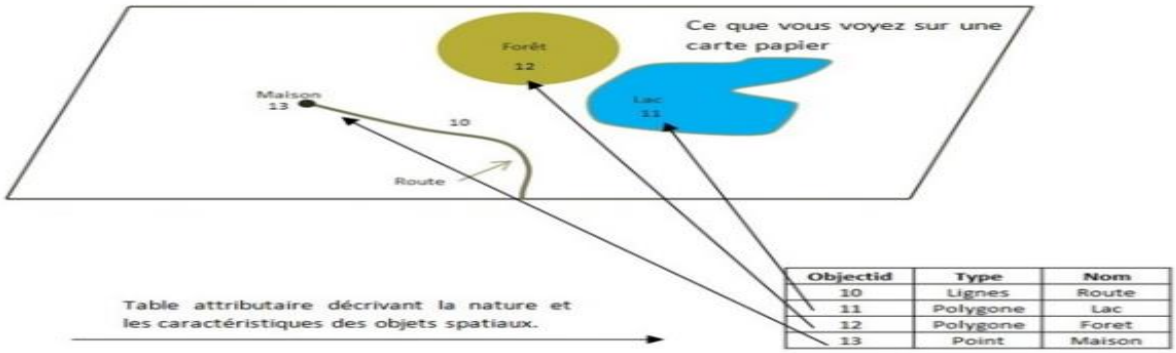

Figure II.6: Exemple des données attributaires dans un SIG

6. Modes de données dans les SIG

Il est d'utiliser différentes techniques pour récupérer des objets cartographiques existants sur support papier afin de les incorporer dans un SIG, telles que la digitalisation et le balayage électronique. La première entraîne directement l'obtention de données cartographiques numériques de type vecteurs, tandis que la seconde entraîne l'acquisition de données raster.

6.1. Les données Vectoriel (mode vecteur)

Les éléments constitutifs du monde réel sont reproduits à travers des points, des lignes (arcs) et des polygones (Fig.2.7). La création de la donnée graphique (objet spatial) va créer automatiquement une ligne dans une table attributaire et donc crée de la donnée alphanumérique, les deux données (alphanumérique et graphique) sont interactives.

- **Les points** : sont utilisés pour représenter des phénomènes ou des objets ponctuels ou d'extension spatiale très réduite.

- **Les lignes** : sont utilisées pour représenter les objets linéaires tels le réseau routier, réseau d'assainissement, réseau hydrographique …, des objets qui ont donc une longueur mais à surface très réduite voire inexistante tel est le cas des courbes de niveau

- **Les polygones** : sont utilisés pour représenter des objets à surface. Un polygone représentera forcément un ou plusieurs objets homogènes tels : les niveaux de risques, modes d'occupation du sol.

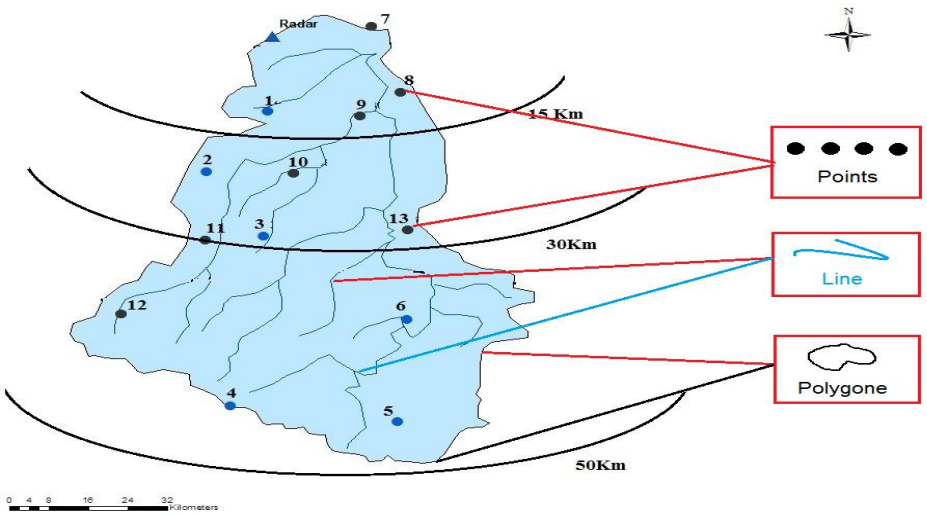

Figure II.7: Exemple d'un mode vecteur

6.1.1. Caractéristique du modèle vectoriel

- Distribution irrégulière des entités spatiales
- Représentation géométrique des entités
- Frontières explicites
- Localisation précise et unique
- Attributs et entités reliés par un numéro d'identification
- Représentation très efficace de la topologie

6.1.2. Quelques formats vectoriels

- DLG (Digital Line Graphs)
- Utilisé pour les cartes topographiques (USGS)
- DXF (Data Exchange Format), format d'Import/Export d'Autocad
- E00 (ungenerate), format d'Import/Export d'Arc/Info
- MID/MIF (MapInfoInterchange File), format d'Import/Export de MapInfo
- Shape file Shp, format spécifique d'Arcgis.

6.2. Les données matricielles (mode raster)

La donnée raster représente le territoire à l'aide d'une grille régulière de pixels, où chaque pixel est défini par sa résolution, sa position, et son intensité de gris (ou couleur). La juxtaposition de ces pixels recrée l'apparence du territoire, avec une taille de pixel plus grande entraînant une moindre information géographique, tandis qu'une taille réduite offre une information de haute résolution, chaque pixel ayant une valeur de teinte et de couleur. Elle décrit la surface cartographique point par point, contrairement au mode vecteur qui ne se

concentre que sur les contours. Il est principalement utilisé dans les systèmes à balayage tels que les scanners et les capteurs de télédétection.

6.2.1. Quelques formats matriciels

Chaque pixel contient une information numérique codée en bits:1, 8, 24,32…

Image en noir et blanc (1bit, 2 couleurs)

-image en niveaux de gris (8 bits, 256 couleurs)

-Standards: BMP, TIFF, JPG, PNG...

-Compressés: ECW, JP2, SID

-Logiciels de SIG et télédétection : GRID, BIL, BSQ, BIP, IMG(ERDAS)...

Exemple source de donné raster

Des photos aériennes, image satellite ou image radar et les documents scannés

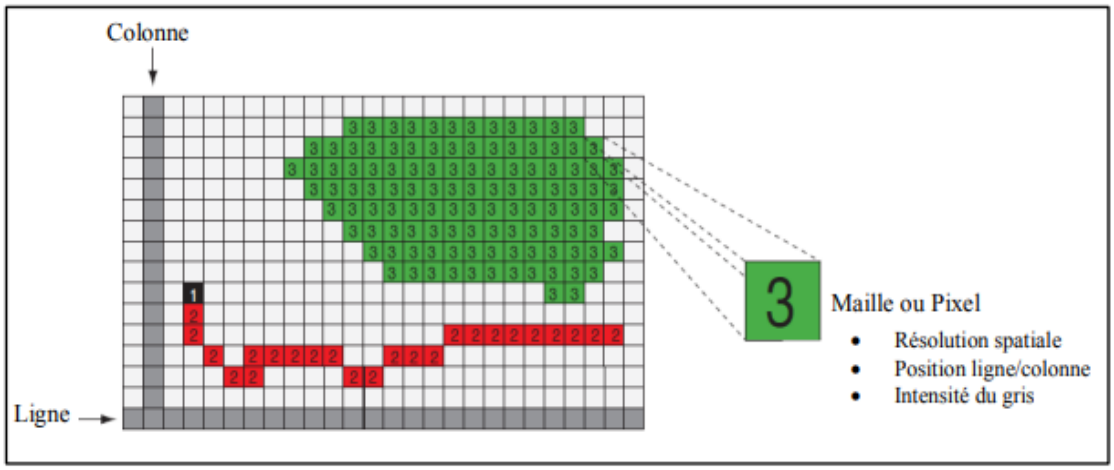

Figure II.8: Exemple d'un mode raster

Conclusion

Ces deux modes se complètent mutuellement. Le raster convient mieux à différentes applications (télédétection) et offre une solution économique à certains besoins. Par exemple, l'administrateur d'un réseau pourrait se limiter à scanner des fonds de plans en les gardant au format raster et en les numérisant ensuite en mode vecteur (ce qui requiert une définition par formes géométriques). Le vecteur englobe tous les besoins fréquents en matière de gestion de données localisées.