

Systemes d'information géographique

Avec les logiciels libres GRASS et QGIS

2^e ÉDITION

Yves Auda

Ingénieur CNRS

Géosciences Environnement Toulouse

Observatoire Midi-Pyrénées

Centre national de la recherche scientifique

Graphisme de couverture : Elizabeth Riba
Illustration de couverture : © M. Svetlana - Shutterstock.com

| | |
|--|--|
| <p>Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.</p> <p>Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements</p> | <p>d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.</p> <p>Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).</p> |
|  | |

© Dunod, 2018, 2022
11, rue Paul Bert, 92240 Malakoff
www.dunod.com
ISBN 978-2-10-083401-3

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

| | |
|--|-----------|
| Introduction | VII |
| Partie 1 | |
| Partie théorique | 1 |
| 1 Qu'est-ce qu'un SIG ? | 2 |
| 1. Définition | 2 |
| 2. Brève histoire | 2 |
| 3. Besoins et utilité | 4 |
| 4. Objet géographique | 4 |
| 5. Vecteur versus Raster | 5 |
| 6. Structure en couches | 7 |
| 7. Système de coordonnées de référence | 8 |
| 8. Briques du SIG | 8 |
| 2 Modèles de données | 10 |
| 1. Modèle Vecteur | 10 |
| 2. Géométrie des vecteurs | 10 |
| 3. Topologie des vecteurs | 11 |
| 4. Opérations topologiques | 14 |
| 5. Requêtes spatiales Vecteur | 17 |
| 6. Modèle Raster | 21 |
| 7. Internet et SIG | 23 |
| 3 Localisation sur Terre | 26 |
| 1. Représentation du globe terrestre | 26 |
| 2. Système géodésique | 27 |
| 3. Changement de système géodésique | 30 |
| 4. Coordonnées géographiques | 31 |
| 5. Coordonnées cartographiques | 33 |

| | |
|--|----|
| 6. Exemples de projection | 34 |
| 7. Changement de système de coordonnées de référence | 39 |
| 8. Calcul d'une distance ou d'une surface | 42 |
| 4 Composition de carte | 44 |
| 1. Spécificité d'une carte géographique | 44 |
| 2. Couleur | 47 |
| 3. Niveaux de gris | 49 |
| 4. Sémiologie graphique des objets | 49 |
| 5. Standardisation des symbologies | 54 |
| 6. Représentation des rasters | 55 |
| 5 Modèles hydrologiques et MNT | 57 |
| 1. Le relief | 57 |
| 2. MNT versus MNE | 58 |
| 3. Formes de MNT | 58 |
| 4. Construction des MNT | 59 |
| 5. Pente et orientation des MNT | 62 |
| 6. Flux hydriques | 66 |
| 7. Bassin versant | 67 |
| 6 Modèles de paysage | 69 |
| 1. Concept de paysage | 69 |
| 2. Métriques paysagères | 70 |
| 3. Distributions statistiques de métriques | 73 |
| 4. Protocoles d'échantillonnage | 74 |
| 5. Implémentations | 76 |
| 7 Traitement des données | 77 |
| 1. Typologie des variables | 77 |
| 2. Emporte-pièce | 79 |
| 3. Précautions d'usage | 79 |
| 4. Exemples types | 80 |

Partie 2

| | |
|--|-----|
| Partie pratique | 85 |
| Exercice 1 Prise en main de GRASS et QGIS | 86 |
| Exercice 2 Repères cartographiques | 93 |
| Exercice 3 Géoréférencement d'une carte scannée | 99 |
| Exercice 4 Organisation des données et de leur traitement | 105 |
| Exercice 5 Base de données et SQL | 115 |
| Exercice 6 SQL spatial, requêtes géométriques | 124 |
| Exercice 7 SQL spatial, requêtes topologiques | 133 |
| Exercice 8 Modèle numérique de terrain | 141 |
| Exercice 9 Les interfaces R du CRAN et Python | 146 |
| Exercice 10 Modèle hydrologique | 156 |
| Exercice 11 Traitement des données | 162 |
| Exercice 12 Modèle de paysage | 175 |
| Exercice 13 Composition de carte | 183 |
| Conclusion | 190 |
| Bibliographie | 191 |
| Sigles et acronymes | 196 |
| Remerciements | 198 |
| Index | 199 |

À la découverte de votre livre

1 Ouverture de chapitre

Elle donne :

- une introduction aux sujets et aux problématiques abordés dans le chapitre
- un rappel des objectifs pédagogiques
- le plan du chapitre

2 Le cours

Le cours, concis et structuré, est structuré en deux parties :

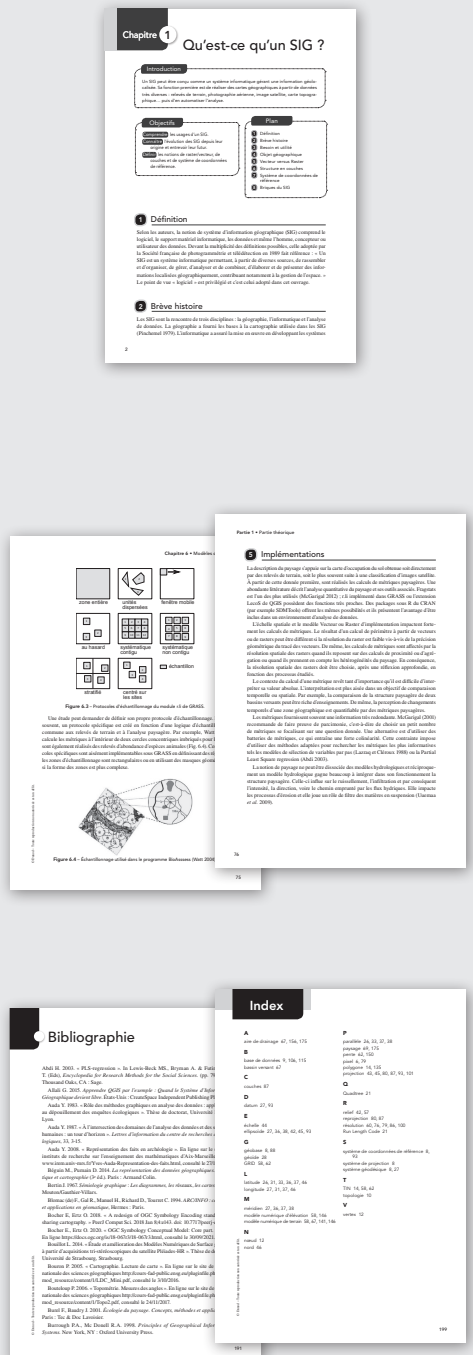
- une partie théorique exposant les notions indispensables à l'utilisation des SIG.
- une partie pratique regroupant des exercices reprenant chaque notion théorique.

3 En fin d'ouvrage

- Une bibliographie
- Une liste des sigles et acronyme
- Un index

4 Ressources numériques

Retrouvez sur www.dunod.com l'ensemble des données liées aux exercices.



Introduction

Cet ouvrage est un recueil de cours et de travaux pratiques dispensés dans le cadre du Master 2 EEA, spécialité « Signal Imagerie, et Applications » de l'université Paul Sabatier de Toulouse. Plusieurs formations données auprès des chercheurs de l'Observatoire Midi-Pyrénées dans le cadre de l'atelier permanent : « Méthodes numériques, données, traitement du signal » ont également permis d'améliorer les supports pédagogiques dans un autre contexte, celui de la formation continue.

Il est composé de deux parties. La partie théorique expose de manière simplifiée les fondements théoriques indispensables à l'utilisation des systèmes d'information géographique. La partie pratique met en œuvre les notions abordées précédemment à partir de données recueillies dans le cadre du Service national d'observation MSEC (Multi-Scale Environmental Changes) dans un bassin versant Houay Xon situé au nord du Laos. Ce choix donne une cohérence thématique aux exemples traités et facilite la compréhension globale des questions environnementales abordées.

Le contenu de cet ouvrage reste valide quel que soit le logiciel utilisé. L'expérience de l'auteur, en particulier dans le monde du logiciel libre, conduit à utiliser principalement GRASS et QGIS pour les exercices. D'autres programmes (R du CRAN, sqlite3...) sont également exploités. Pour faire partie de notre logithèque, un programme doit être diffusé sous une licence « open source » et il doit exister une version pour chacun des trois systèmes d'exploitation Linux, Windows et MacOS. Ce choix de privilégier dans l'enseignement l'usage des logiciels libres est préconisé par l'article L. 123-4-1 de la loi ESR du 9 juillet 2013 du parlement français.

Partie 1

Partie théorique

Le premier chapitre présente les principes du SIG et donne une vue d'ensemble de chacun de ses composants. Il explique les étapes qui ont conduit à la construction des SIG sous leur forme actuelle puis il introduit les notions d'objet géographique, de type de données raster et vecteur, de couches et de projection qui sont approfondies dans les chapitres suivants. Le chapitre 2 décrit l'organisation des données dans un SIG en s'appuyant sur les standards établis par l'OGC. La topologie des vecteurs et la structure des rasters sont décrites en attachant une attention particulière à la performance du stockage et des traitements. Le chapitre 3 résume les principes de localisation sur la terre par une description simplifiée des systèmes géodésiques et des projections, ces deux éléments étant regroupés sous le terme de système de coordonnées de référence. Il insiste sur les erreurs à ne pas commettre lors d'un changement de SCR. Le chapitre 4 expose les notions utiles à la réalisation d'une carte géographique. Il allie concepts cartographiques et sémiologie graphique qui ne peuvent être dissociés dans la réalisation d'une carte. Le chapitre 5 est dédié aux modèles numériques de terrain et leur implication dans les processus hydrologiques. Le chapitre 6 aborde les modèles de paysage. Les métriques paysagères sont présentées en insistant sur l'influence de la précision géométrique des données. Le chapitre 7 rend compte d'une expérience de l'auteur dans l'analyse des données rasters. Il propose un mode d'organisation des données qui sera utilisée dans la partie pratique. L'objectif de cette partie théorique n'est pas de réaliser un état de l'art complet de ces sujets mais plutôt de fournir les fondements utiles à la compréhension des exercices.

Qu'est-ce qu'un SIG ?

Introduction

Un SIG peut être conçu comme un système informatique gérant une information géolocalisée. Sa fonction première est de réaliser des cartes géographiques à partir de données très diverses : relevés de terrain, photographie aérienne, image satellite, carte topographique... puis d'en automatiser l'analyse.

Objectifs

Comprendre les usages d'un SIG.

Connaître l'évolution des SIG depuis leur origine et entrevoir leur futur.

Définir les notions de raster/vecteur, de couches et de système de coordonnées de référence.

Plan

- 1 Définition
- 2 Brève histoire
- 3 Besoin et utilité
- 4 Objet géographique
- 5 Vecteur versus Raster
- 6 Structure en couches
- 7 Système de coordonnées de référence
- 8 Briques du SIG

1 Définition

Selon les auteurs, la notion de système d'information géographique (SIG) comprend le logiciel, le support matériel informatique, les données et même l'homme, concepteur ou utilisateur des données. Devant la multiplicité des définitions possibles, celle adoptée par la Société française de photogrammétrie et télédétection en 1989 fait référence : « Un SIG est un système informatique permettant, à partir de diverses sources, de rassembler et d'organiser, de gérer, d'analyser et de combiner, d'élaborer et de présenter des informations localisées géographiquement, contribuant notamment à la gestion de l'espace. » Le point de vue « logiciel » est privilégié et c'est celui adopté dans cet ouvrage.

2 Brève histoire

Les SIG sont la rencontre de trois disciplines : la géographie, l'informatique et l'analyse de données. La géographie a fourni les bases à la cartographie utilisée dans les SIG (Pinchemel 1979). L'informatique a assuré la mise en œuvre en développant les systèmes

de base de données pour stocker les données, les interfaces homme machine (IHM) pour le dialogue et toutes les procédures de traitements relevant de l'analyse de données. Dans ce domaine, Bertin (1967) a initié les premiers traitements graphiques de données spatialisées. Ils ont constitué les prémices à une automatisation rendue faisable par l'évolution technologique des ordinateurs. Rimbart (1990) a décrit ces possibilités offertes par l'alliance de la géographie et des traitements automatisés que permet l'informatique.

Les SIG sont nés de l'association de la cartographie et de l'informatique. Il est cependant difficile de tracer l'origine des SIG, d'identifier le moment où ces trois mondes de la géographie, de l'informatique et de l'analyse de données se sont rejoints. Le premier exemple cité date de 1854 quand le docteur John Snow a représenté la localisation des malades atteints par le choléra dans le quartier de Soho à Londres, en Angleterre. Mais ce n'est qu'à partir des années 1960 que l'informatique a permis l'essor des SIG. Le premier SIG aurait été développé par Roger Tomlinson pour inventorier les terres du Canada dans les années 1960. L'objectif était de réhabiliter les terres agricoles. D'autres initiatives reposant sur des développements spécifiques apparurent à la même époque aux États-Unis. Le département de l'eau et de contrôle de la pollution créa un système d'acquisition et d'extraction de la qualité des eaux. Le service forestier conçut un système de gestion des parcelles qui conduisit à la production de cartes à partir de couches d'information (de Blomac *et al.* 1994). Puis de 1960 à 1970, virent le jour des produits non spécifiques à un jeu de données : SYMAP, CALFORM, SYMVU, GRID, POLYVRT, ODYSSEY développés dans le Hardward Lab for Computer Graphics (Malczewski 2004).

Mais les SIG ont vraiment pris leur essor avec le développement des systèmes de base de données. Le plus populaire, ARCINFO, reposant sur la base de données PRIME, est commercialisé en 1982 par la société ESRI (Environmental Systems Research Institute). À la même époque, les ingénieurs de l'armée américaine et les chercheurs de l'université de l'Illinois commencèrent à développer un SIG qui devint, quelques années plus tard, GRASS (Westervelt 2004). Dès le début, sa philosophie de développement a été axée sur l'ouverture du code offrant la possibilité à tout programmeur de participer à son écriture. L'interface était minimaliste et ce n'est que dans les années 1990 que GRASS a constitué une alternative aux SIG commerciaux. L'avènement de la micro-informatique a ensuite conduit à la création de nombreux produits dont il est difficile d'établir une liste exhaustive : IDRISI, MapInfo, ENVI, SAVANE, ILWIS, APIC, GeoConcept, SPRING, GvSIG, QGIS... L'adoption de standard pour la définition des systèmes de projection (proj.4) et pour le format de fichiers géographiques (GDAL) supportés par la fondation OSGeo ont assuré l'échange de données entre ces SIG.

Le développement des services Internet dans les années 1990 a donné accès à de nouvelles sources de données. Sous l'égide de l'Open Geospatial Consortium (OGC), différents standards de formats ouverts de données, de protocoles... garantissent l'interopérabilité et favorisent l'accès à des bases de données spatiales souvent libres d'accès. Citons Web Map Service (WMS), Web Map Tile Service (WMTS), Web Feature Service (WFS), Web Coverage Service (WCS), Catalogue Service for the Web (CS-W), Web Processing Service (WPS), Simple Feature Access (SFS), Geography Markup Language

(GML), Keyhole Markup Language (KML), Sensor Web Enablement (SWE), Table Joining Service (TJS). Dans ce contexte, QGIS créé en 2002 s'est imposé. La raison de son succès est liée à la facilité d'écrire de nouvelles extensions.

Très récemment, le Web sémantique a inclus des ontologies adaptées à l'information spatiale et temporelle. Ce domaine fait actuellement l'objet de recherches pour formaliser la représentation des connaissances et le raisonnement (Feyaerts 2010).

3 Besoins et utilité

Le traitement d'une information spatialisée est grandement facilité par les SIG. La nécessité est d'autant plus grande que le volume des données est important et que l'organisation de ces données est complexe. De nombreux domaines traitent d'une information spatialisée : l'aménagement du territoire, l'étude d'impact d'une construction, la gestion des réseaux, l'agriculture de précision, le suivi en temps réel des véhicules, les risques naturels, la protection civile, la prospection minière, les études de démographie, la socioéconomie et la santé, le géomarketing, etc. Cette liste n'est naturellement pas exhaustive.

Les SIG sont adoptés par de nombreuses institutions, des entreprises, des organismes de recherche, des écoles, des administrations pour des besoins de gestion, de formation ou de recherche.

Les résultats obtenus sont multiples. La recherche du cheminement le plus économique en termes de temps pour la distribution du courrier par le facteur, du tracé dont l'impact sera le plus faible pour l'implantation d'une nouvelle ligne électrique, de la localisation du central téléphonique optimisant le raccord des abonnés sont quelques exemples souvent cités. Mais l'application la plus courante est la réalisation de cartes élaborées suite à des traitements de données parfois complexes. Des données issues de relevés terrain, de photographies aériennes, d'images de télédétection sont utilisées pour l'ensemble des applications précitées. Leur nature très différente exige pour les traiter de les concevoir comme des objets géographiques.

4 Objet géographique

Tout objet géographique possède deux caractéristiques : sa localisation et ses attributs (Fig. 1.1). Par exemple, un arbre est situé à un endroit défini par sa position exprimée en latitude/longitude. Ses attributs le caractérisent par son espèce, sa hauteur, etc. Le SIG doit assurer le stockage des données correspondantes et permettre des requêtes sur les valeurs prises par ces deux caractéristiques. Cette gestion dépend du modèle de représentation des objets géographiques.

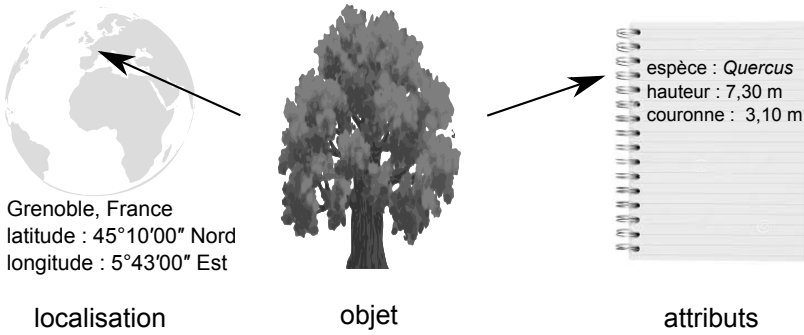


Figure 1.1 – L'objet géographique.

5 Vecteur versus Raster

Toute démarche scientifique impose de créer des modèles pour appréhender la réalité, de facto à tout jamais inaccessible, car tout objet ne peut être perçu qu'au travers d'un modèle (Auda 2008). Les SIG obéissent à cette logique en représentant le monde par deux modèles : le modèle Vecteur et le modèle Raster (Fig. 1.2).

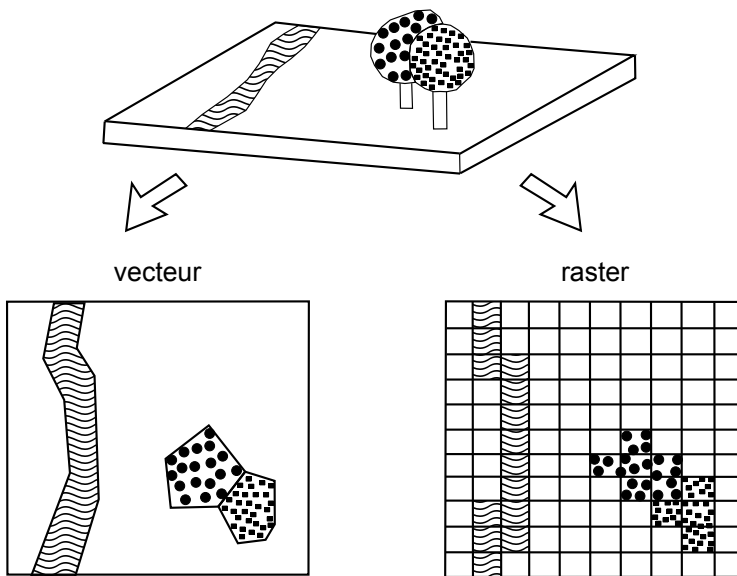


Figure 1.2 – Modèle Vecteur versus Raster.

- Dans le **modèle Vecteur**, les objets sont modélisés par des éléments géométriques. Les objets ponctuels sont représentés par un simple point. Les objets linéaires (route, fleuve...) sont constitués de lignes. Les objets surfaciques (territoire géographique,

parcelle...) sont résumés sous forme de polygones. Les propriétés de ces objets sont stockées dans une table attributive annexe.

- Dans le **modèle Raster**, les objets sont modélisés par une grille dont les éléments sont appelés pixels. La valeur de chaque pixel caractérise la partie de l'objet correspondante. La localisation est simplement définie par la position du coin supérieur gauche de l'image et la taille du pixel.

Ces deux modèles Vecteur/Raster impliquent des structures de stockage des données très différentes (Fig. 1.3).

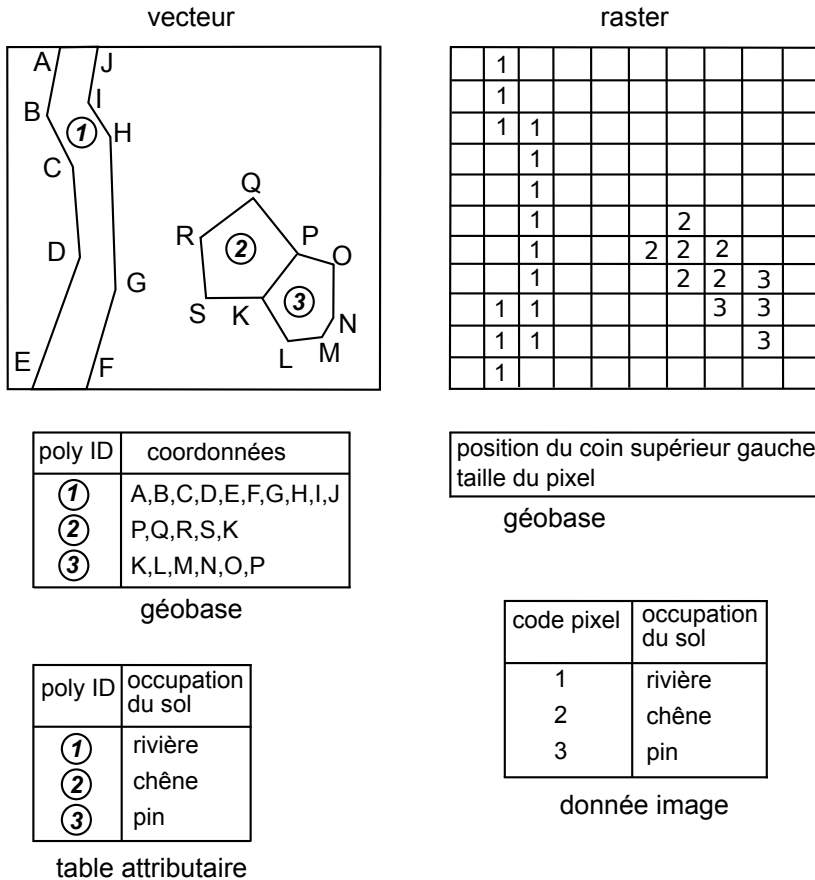


Figure 1.3 – Stockage des données dans les modèles Vecteur et Raster.

Au tout début de l'histoire des SIG, rares étaient les logiciels qui permettaient de gérer les deux modèles. Un SIG était orienté Vecteur (ArcInfo, MapInfo, APIC...) ou Raster (IDRISI, ENVI, ERDAS...). Aujourd'hui, GRASS et QGIS permettent la conversion de données d'un modèle à l'autre et le stockage de ces deux modèles dans une même base de données, les vecteurs et les rasters constituant des couches séparées.