

Les systèmes d'information géographique

Dans la même collection

- BAUELLE Guy, *Géographie du peuplement*, 2016, 3^e édition.
- BÉGUIN Michèle, PUMAIN Denise, *La représentation des données géographiques. Statistique et cartographie*, 2017, 4^e édition.
- BOULAY Guilhem, GRANCLEMENT Antoine, *Introduction à la géographie économique*, 2019.
- CIATTONI Annette, VEYRET Yvette (dir.), *Les fondamentaux de la géographie*, 2018, 4^e édition.
- DAVID Olivier, *La population mondiale. Répartition, dynamique et mobilité*, 2015, 3^e édition.
- DESJARDINS Xavier, *L'aménagement du territoire*, 2017.
- DI MÉO Guy, *Introduction à la géographie sociale*, 2014.
- FEUILLET Thierry, COSSART Étienne, COMMENGES Hadrien, *Manuel de géographie quantitative*, 2019.
- GÆURY David, SIERRA Philippe, *Introduction à l'analyse des territoires*, 2016.
- GRATALOUP Christian, *Introduction à la géohistoire*, 2015.
- GUINARD Pauline, *Géographies culturelles*, 2019.
- HUMAIN-LAMOURE Anne-Lise, LAPORTE Antoine, *Introduction à la géographie urbaine*, 2017.
- LOUCHET André, *Les océans. Bilan et perspectives*, 2013.
- LOUISET Odette, *Introduction à la ville*, 2011.
- LOYER Barbara, *Géopolitique. Méthodes et concepts*, 2019.
- MORANGE Marianne, SCHMOLL Camille, *Les outils qualitatifs en géographie*, 2016.
- PUMAIN Denise, SAINT-JULIEN Thérèse, *Analyse spatiale. Les interactions*, 2010, 2^e édition.
- PUMAIN Denise, SAINT-JULIEN Thérèse, *Analyse spatiale. Les localisations*, 2010, 2^e édition.
- REGHEZZA-ZITTA Magali, *La France dans ses territoires*, 2017, 2^e édition.
- VEYRET Yvette, LAGANIER Richard, SCARWELL Helga-Jane, *L'environnement. Concepts, enjeux, territoires*, 2017.

CHRISTINA ASCHAN-LEYGONIE
CLAIRE CUNTY
PAULE-ANNICK DAVOINE

Les systèmes d'information géographique

Principes, concepts
et méthodes

ARMAND COLIN

Conception de couverture : Hokus Pokus créations
Image de couverture : Christina Aschan-Leygonie, Claire Cunty,
Paule-Annick Davoine ; Source : BD Alti®/IGN
Maquette intérieure : Raphaël Lefeuvre
Mise en page : Belle page
Réalisation des figures : Cyrille Suss et Claire Cunty

<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>
	

© Armand Colin, 2019

Armand Colin est une marque de
Dunod Éditeur, 11 rue Paul Bert, 92240 Malakoff
ISBN : 978-2-200-61718-9

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.

Sommaire

Introduction	9
1. De la « réalité » du terrain à l'information géographique dans les SIG	13
1. Construire un modèle spatial de la « réalité » du terrain	14
1.1 Des données à l'information géographique	15
1.2 Phénomènes géographiques continus, discrets, en réseau	16
1.3 La traduction du monde réel par une sélection de l'information géographique	18
1.4 La notion de couche d'information géographique thématique	18
1.5 La modélisation numérique de l'information géographique	20
2. Deux modes de représentation en deux dimensions	21
2.1 Le mode raster	23
2.2 Le mode vecteur	25
3. La structuration de l'information géographique	32
3.1 Le modèle vecteur non topologique	34
3.2 Le modèle vecteur topologique	36
3.3 Le réseau, un modèle vecteur particulier intégrant la topologie	38
2. Fondamentaux de l'information géographique	43
1. Les systèmes de coordonnées	44
1.1 Les systèmes de coordonnées géographiques	45
1.2 Les systèmes de coordonnées projetées	51
1.3 Les systèmes de coordonnées dans les SIG	57
2. La troisième dimension de l'information géographique	60
2.1 La troisième dimension dans les jeux de données géographiques	61
2.2 La modélisation de surfaces de terrain	62
2.3 La modélisation d'objets discrets avec une géométrie tridimensionnelle	68

3. Les relations des objets dans l'espace : la topologie	70
3.1 Les relations topologiques	70
3.2 Les opérateurs topologiques	75
4. La distance entre les objets géographiques : une notion fondamentale	77
4.1 Les types de distances dans les SIG	79
4.2 La prise en compte des relations de distances dans les analyses	91
3. Intégration de données géographiques dans les SIG	98
1. Mobiliser des jeux de données géographiques existants	100
1.1 Les sources et fournisseurs d'information géographique	100
1.2 Les données ouvertes ou l' <i>Open Data</i>	102
1.3 Les services Web géographiques pour accéder à l'information géographique	104
1.4 Les métadonnées pour caractériser, comprendre et échanger l'information géographique	106
2. Intégrer et créer des jeux de données géographiques dans un SIG	110
2.1 Intégrer l'information géographique en mode vecteur	110
2.2 Intégrer l'information géographique en mode raster	120
3. Convertir des jeux de données	126
3.1 La rasterisation	126
3.2 La vectorisation	128
4. Stocker des jeux de données géographiques : les formats informatiques	129
4.1 Stocker des jeux de données raster	129
4.2 Stocker des jeux de données vecteur	135
5. SIG et bases de données	140
5.1 Les systèmes de gestion de bases de données relationnels (SGBD)	141
5.2 Les bases de données géographiques	144
5.3 Connexions entre SIG et bases de données relationnelles	146

4. Visualisation, interrogation et enrichissement de l'information géographique	149
1. Visualiser l'information géographique	150
1.1 Organiser les données dans un document cartographique	150
1.2 Naviguer à travers les données géographiques – les jeux de données à l'écran	152
1.3 Personnaliser l'affichage	153
2. Enrichir sémantiquement l'information géographique	156
2.1 Enrichir des jeux de données vecteur	156
2.2 Enrichir des jeux de données raster	164
3. Interroger et sélectionner l'information géographique	166
3.1 Les principes généraux d'interrogation des données	166
3.2 Les critères d'interrogation des données	168
4. Communiquer l'information géographique	174
4.1 La sélection de l'information à représenter	175
4.2 La généralisation des tracés	176
4.3 La cartographie de l'information thématique	177
4.4 La cartographie lissée	180
4.5 Les affichages graphiques et les résumés statistiques	182
4.6 La construction d'un document cartographique	182
4.7 La gestion des étiquettes	183
5. Traitements spatiaux de l'information géographique	187
1. Agréger des entités d'un même jeu de données géographiques	188
1.1 Pourquoi agréger des entités ou des cellules ?	189
1.2 Agréger spatialement des entités d'un jeu de données vecteur	190
1.3 Agréger spatialement des cellules d'un jeu de données raster	193
2. Délimiter une partie d'un jeu de données géographiques à partir d'entités	194
2.1 Extraire/découper	195
2.2 Créer des zones tampons	197
2.3 Créer une enveloppe convexe ou concave autour d'entités vecteur	200

3. Combiner les jeux de données géographiques	203
3.1 Combiner la géométrie de jeux de données vecteur	203
3.2 Combiner des jeux de données raster	214
3.3 Combiner des jeux de données raster et vecteur	215
4. Le schéma de traitements	216
6. Aperçu de l'analyse spatiale avec les SIG	222
1. Analyses intégrant la distance	224
1.1 Aires d'attraction	225
1.2 Autocorrélation spatiale, analyses de concentration et interpolation spatiale	233
2. Analyses de surface	241
2.1 Les analyses de pente, de profil, d'exposition et de volume	241
2.2 Les analyses de visibilité	243
2.3 Les analyses hydrologiques	245
3. Analyse spatiale multicritère	246
3.1 Démarche d'analyse spatiale multicritère dans un SIG	249
3.2 Deux types de critères : « contraintes » et « facteurs »	250
Bibliographie	257
Table des sigles	259
Index	261
Remerciements	269

Introduction

L'information géographique prend aujourd'hui une ampleur croissante, aussi bien en termes de volume que de disponibilité. Elle est au cœur de la géographie, l'aménagement, l'environnement, l'urbanisme, mais s'étend à de nombreux autres domaines ou activités économiques. La production de données géographiques s'est largement démocratisée en raison notamment des dispositifs mobiles dotés de systèmes de géolocalisation. La dimension spatiale s'invite donc dans la description et l'analyse de nombreux phénomènes, qu'ils soient environnementaux, territoriaux, économiques, sociétaux, historiques, etc. Elle se concrétise à travers des représentations cartographiques, comme en témoignent les nombreuses productions du Web. Les finalités de la carte sont aujourd'hui largement reconnues : localisation d'objets, de phénomènes, voire d'information ; mise en évidence des phénomènes spatialisés et révélation du fonctionnement de systèmes spatiaux ; contribution à la prise de décision territoriale... Cependant, la construction cartographique passe au préalable par des étapes de modélisation, d'acquisition, de transformation, de combinaison, de traitements statistiques et d'analyse de l'information géographique, qui peuvent être plus ou moins complexes selon les objectifs fixés. C'est notamment pour faciliter toutes ces étapes, et plus largement pour aider à la gestion des territoires, qu'ont été développés à partir des années 1960 les SIG (systèmes d'information géographique).

D'un point de vue fonctionnel, les SIG permettent l'acquisition, le stockage, l'interrogation, le traitement, l'analyse et la visualisation de données géographiques. Ils se distinguent des autres systèmes d'information par leur capacité à gérer des informations relatives à des objets ou des phénomènes localisés à la surface de la Terre ; ils permettent de traiter une série de questions liées à la localisation (Que trouve-t-on à tel endroit ? Où se situe tel phénomène ?), à la répartition spatiale des phénomènes et aux liens et interactions spatiales qui peuvent exister entre eux (À quelle distance est situé le lieu le plus proche ?), à leur

évolution dans l'espace en fonction du temps (Où se situent les changements ? À quels moments se sont-ils produits ?), etc.

Au fil des années, les SIG se sont imposés comme des outils indispensables à la compréhension et la gestion des territoires. Ils sont devenus plus facilement accessibles avec le développement de logiciels SIG bureautiques commerciaux et la montée en puissance de logiciels et librairies informatiques libres de gestion et de traitement de l'information géographique. Souvent simplement vus comme des outils logiciels, ils doivent être envisagés dans une dimension plus large de projet, dans lequel l'information géographique est mobilisée, produite, traitée et analysée avec des méthodes spécifiques pour répondre à un questionnement ayant une dimension spatiale. Ces projets peuvent concerner aussi bien des productions cartographiques (type atlas cartographiques, cartes de synthèse et analytiques, création de documents d'urbanisme ou de planification territoriale), que la gestion et le partage de données dans des contextes variés (gestion des risques, urbanisme opérationnel, suivis environnementaux). Dans ces projets se pose toujours la question des caractéristiques de l'information géographique, de son traitement et de son analyse pour répondre aux objectifs.

Cet ouvrage pédagogique porte sur les principes, concepts et méthodes communs à différents contextes (institutionnels, logiciels, disciplinaires...) de mise en œuvre de projets SIG. Il est le fruit de nombreuses années d'enseignement sur l'information géographique numérique, les SIG, les systèmes de gestion de bases de données, l'analyse spatiale, la cartographie et les statistiques, essentiellement pour des étudiants en géographie et aménagement et en école d'ingénieurs. Nous avons la conviction que la connaissance et la maîtrise des SIG ne doivent pas être subordonnées à l'apprentissage d'un logiciel, qui n'est qu'un aspect des SIG. Les logiciels SIG sont aujourd'hui très performants, mais parfois sous-exploités, notamment par manque de formation théorique et méthodologique. La formation sur les SIG s'effectue généralement au travers de l'apprentissage d'un logiciel spécifique, comme en témoigne par exemple la diversité des ressources pédagogiques disponibles sur le Web qui leur sont dédiées. Ce type d'apprentissage permet aux étudiants certes de produire des résultats, mais sans toujours connaître les tenants et les aboutissants des méthodes mobilisées pour les obtenir. Cette approche est peu propice à un usage efficace de l'information géographique, à une bonne connaissance des potentialités des SIG ou encore à l'adaptation de l'utilisateur aux évolutions technologiques offertes par la géomatique.

L'objectif de l'ouvrage n'est donc pas de former le lecteur à l'utilisation d'un logiciel dédié, mais de lui permettre de s'appropriier les principes des SIG, d'acquérir une connaissance et une compréhension des notions et des méthodes clés. Nous pensons donc que le déploiement informatique d'un projet SIG est relativement aisé à partir du moment où l'utilisateur a une solide connaissance des phénomènes géographiques (natures, structures, interactions), de la façon de les formaliser numériquement, et de la démarche à mettre en œuvre pour résoudre un problème géographique. Par ailleurs, le développement des projets SIG s'appuie sur des concepts, des principes et des méthodes issus des systèmes de bases de données, de l'analyse statistique, de l'analyse spatiale, de la cartographie. Pour utiliser de manière correcte et efficace les SIG, il est nécessaire d'avoir des connaissances dans ces domaines qui ne sont pas présentés de manière approfondie dans cet ouvrage, mais pour lequel le lecteur trouvera au cours du texte des références utiles.

Ce manuel s'adresse à un public non spécialisé en science de l'information géographique (ou géomatique). Il est avant tout écrit pour des étudiants de licence ou master de géographie et aménagement, discipline académique qui a les plus forts liens avec les SIG. Cependant, l'ouvrage s'adresse aussi à d'autres publics : étudiants d'autres disciplines ; enseignants, chercheurs, techniciens, ingénieurs ou tout autre professionnel souhaitant se former aux SIG. Il est organisé en six chapitres et a été structuré pour former un tout unifié, mais chaque chapitre peut être lu séparément.

Le **premier chapitre** porte sur la numérisation de l'information géographique : comment l'information géographique modélise-t-elle la réalité du terrain ? Quelles sont ses composantes et comment sont-elles structurées au sein des SIG ?

Le **deuxième chapitre** traite des caractéristiques fondamentales de l'information géographique. Il aborde les systèmes de coordonnées et la troisième dimension, ainsi que les relations topologiques et de distances liées à la position relative des objets géographiques.

Le **troisième chapitre** approfondit les méthodes d'intégration de données dans les SIG. Celles-ci sont variées : importation ; géocodage ; création et transformation de jeux de données géographiques. Il présente aussi les modalités d'échanges de données dans le contexte de l'*Open Data* et les différents formats informatiques utilisés pour stocker et gérer les jeux de données géographiques.

Le **quatrième chapitre** présente les fonctionnalités premières et fondamentales des SIG, qui sont systématiquement les plus mobilisées lors de la réalisation d'un projet, quelles que soient sa finalité et son ampleur : visualisation, enrichissement sémantique, interrogation des données, réalisation d'un document cartographique.

Le **cinquième chapitre** décrit les principaux traitements spatiaux, appelés aussi géotraitements, proposés dans les SIG. Ces traitements sont très variés et permettent notamment de transformer des jeux de données ou d'en extraire des sous-ensembles en mobilisant des méthodes variées. L'enchaînement de différents traitements dans un projet SIG se traduit par un schéma de traitements.

Le **sixième chapitre** est dédié à une présentation des méthodes d'analyses spatiales les plus fréquemment mobilisées dans un projet SIG. Celles-ci préexistent aux SIG et s'appuient sur des fondements théoriques forts, qui font déjà l'objet d'ouvrages dont nous recommandons la lecture pour comprendre et maîtriser les enjeux géographiques et méthodologiques qui sous-tendent leurs mises en œuvre.

Il n'existe pas véritablement de consensus sur l'ensemble des termes associés aux SIG et nous avons fait des choix de termes qui sont toujours explicités dans le texte. Par ailleurs, les méthodes et fonctionnalités abordées sont systématiquement appliquées aux données en mode vecteur et en mode raster. Enfin, les illustrations sont des éléments fondamentaux de cet ouvrage, elles sont souvent accompagnées d'exemples pour faciliter la compréhension *via* des cas d'application.

Plusieurs éléments pédagogiques et rubriques communs à chaque chapitre sont proposés pour aider le lecteur à tirer le meilleur parti de cet ouvrage :

- Les objectifs d'apprentissage sont précisés au début de chaque chapitre.
- Les erreurs à ne pas commettre mettent en garde sur des éléments qui posent fréquemment des problèmes, comme nous avons pu le constater dans nos expériences d'enseignement.
- Un pas de côté est proposé dans la rubrique « Le saviez-vous ? ».
- Les Focus permettent d'approfondir certaines questions.
- Les notions à maîtriser, classées par thématiques, qui concluent chaque chapitre, permettent de rassembler le vocabulaire spécifique au domaine.
- Des questions pour réviser sont proposées en fin de chapitre.

■ Chapitre 1

De la « réalité » du terrain à l'information géographique dans les SIG

Objectifs

- Comprendre la nécessité de conceptualiser pour numériser le monde réel et le traduire en information géographique.
- Distinguer les phénomènes géographiques continus et discrets.
- Comprendre la différence entre une donnée géographique et une information géographique.
- Connaître les principes de modélisation et de structuration de l'information géographique dans un SIG.
- Comprendre l'intérêt de la structuration topologique des données en mode vecteur.

Trois étapes sont nécessaires pour passer du monde réel à l'implémentation informatique sous la forme d'un modèle de données. La première étape suit une démarche d'abstraction, au cours de laquelle sont identifiés les différents phénomènes géographiques observés sur le terrain que

l'on souhaite représenter dans le SIG pour mener à bien les analyses. La deuxième étape consiste à choisir le modèle de données à travers lequel les données peuvent être stockées, affichées et analysées : c'est la modélisation numérique. Dans la troisième étape, il s'agit de voir comment les relations entre les unités d'observation du phénomène seront structurées.

1. Construire un modèle spatial de la « réalité » du terrain

Le terrain est vu par un observateur qui interprète ce qu'il voit et catégorise l'information, souvent de façon inconsciente. Le passage de la « réalité » (sans rentrer dans une controverse philosophique sur ce qu'est la réalité) à son observation, puis à son enregistrement sous forme numérique pour pouvoir être intégrée dans un SIG, dépend donc du filtre perceptif de l'observateur. Pour choisir l'information à prendre en compte et à traduire sous forme numérique, l'observateur trie ses observations en fonction de ses objectifs, mais aussi en fonction de ses caractéristiques propres (sa culture, ses connaissances...). Aucune donnée ne doit donc être considérée comme totalement neutre et objective. Il est nécessaire d'en avoir conscience pour interpréter correctement et avec prudence les résultats issus des traitements réalisés par les SIG, sans pour autant s'interdire de les utiliser !

On appelle souvent modélisation¹ (voir Focus, p. 20) ce processus d'abstraction de la réalité qui permet de la simplifier pour aider à la comprendre. Ce dernier, pour rester cohérent avec la réalité, a pour objectif de décrire non seulement les objets mais également les relations qui existent entre eux afin de prendre en compte la manière dont l'ensemble des objets se structurent dans l'espace. Ces relations sont liées à leurs positions relatives, décrivant ainsi des propriétés de voisinage, d'inclusion ou d'intersection (voir chapitre 2, p. 70).

1. Le terme de représentation est aussi souvent utilisé, mais nous préférons ne pas l'employer afin de ne pas créer de confusion avec la représentation cartographique au sens de la visualisation des données sur une carte.

1.1 Des données à l'information géographique

Pour comprendre ce que signifie l'information géographique, il est utile de partir du terme de « donnée localisée » (également appelée « donnée géographique » ou « donnée spatiale »). Toute donnée qui se réfère à une localisation, *via* des coordonnées géographiques (voir chapitre 2, p. 43), un nom de lieu ou une adresse, peut être considérée comme une donnée géographique. Les données localisées peuvent être obtenues par une mesure (exemple : bruit), une imagerie satellite (exemple : occupation du sol), un inventaire (exemple : commerces), un recensement (exemple : population), une enquête (exemple : déplacements), etc. Les méthodes d'acquisition des données, la création des données en vue d'une utilisation dans un SIG et les sources de données sont décrites dans le chapitre 3.

Il est fréquent de distinguer donnée et information, bien que les deux termes soient souvent utilisés de façon interchangeable. Nous pouvons considérer que les données recueillies sont transformées en informations géographiques lorsqu'elles sont intégrées dans un contexte et qu'elles prennent alors une signification particulière, qu'elles sont interprétées par rapport à une question donnée.

Illustrons par un exemple la distinction entre donnée et information géographique : les membres d'une équipe de recherche sont mandatés pour proposer des actions pour adapter un territoire au changement climatique. Trois chercheurs relèvent les températures de l'air sous couvert forestier et hors couvert forestier et effectuent des entretiens avec les habitants pour recueillir leurs perceptions de la température dans différents lieux. Les résultats obtenus sont transcrits dans des tableaux ; ce sont les données. Ces données, exprimées sous la forme de nombres ou de texte, n'ont pas de signification particulière en elles-mêmes. Pour les rendre utiles et les transformer en information géographique, il est nécessaire de préciser le contexte de l'étude et les méthodes de recueil des données. Elles doivent également être accompagnées de leurs unités de mesures et de leur localisation. Ces précisions sont généralement fournies dans ce qui est appelé « métadonnées » (voir chapitre 3, p. 106), c'est-à-dire des données décrivant les données.