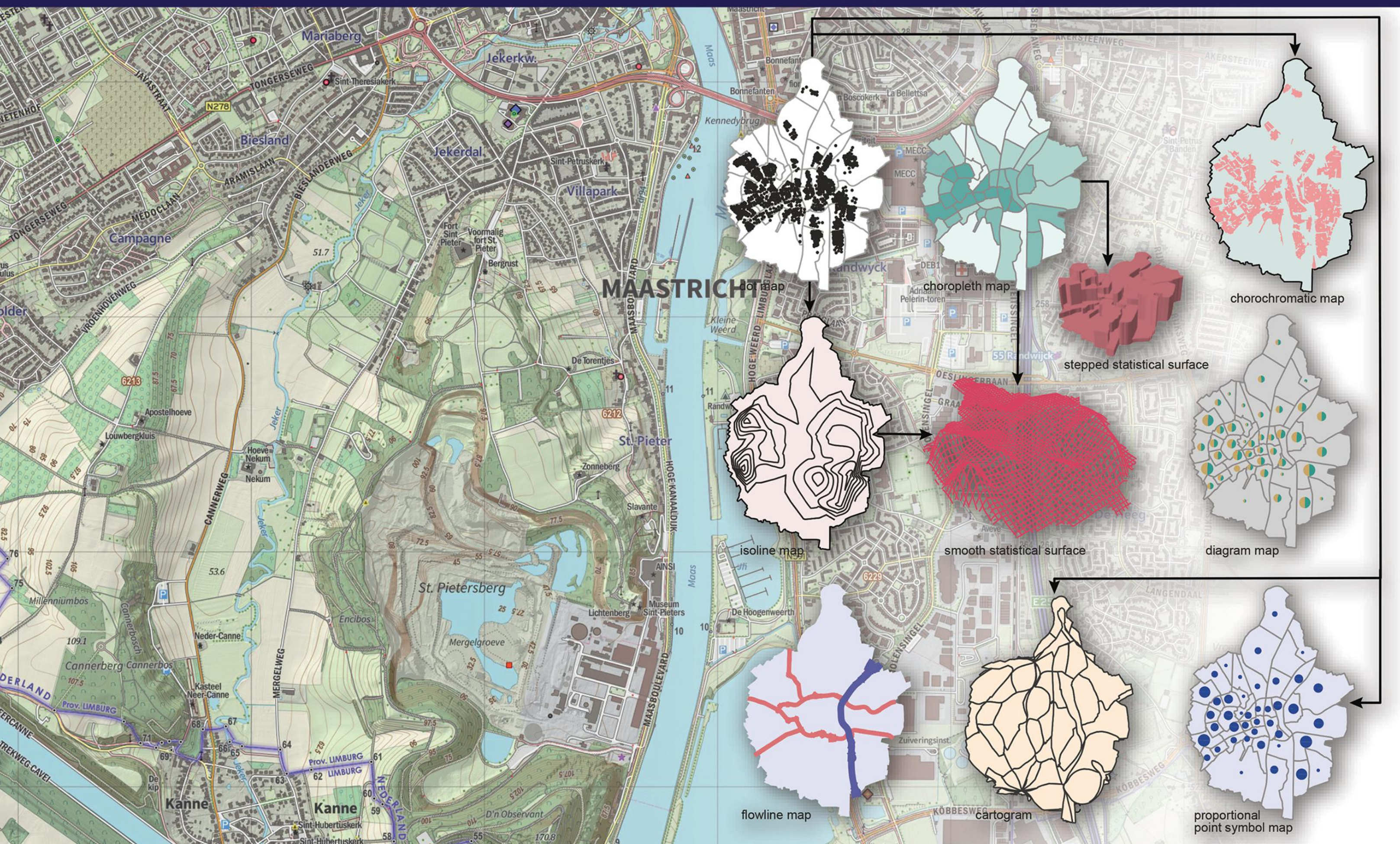
FOURTH EDITION

CARTOGRAPHY

Visualization of Geospatial Data



Menno-Jan Kraak
Ferjan Ormeling

CRC Press
Taylor & Francis Group

Cartography

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Cartography
Visualization of Geospatial Data

Menno-Jan Kraak and Ferjan Ormeling

Fourth Edition

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Fourth edition published 2021
by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

and by CRC Press
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

© 2021 Taylor & Francis Group, LLC

Third edition published by CRC Press 2009

CRC Press is an imprint of Taylor & Francis Group, LLC

Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data
Names: Kraak, M. J., author. | Ormeling, Ferjan, 1942- author.
Title: Cartography : visualization of geospatial data / Menno-Jan Kraak, Ferjan Ormeling.
Description: Fourth edition | Boca Raton ; London : CRC Press, 2020. |
Includes bibliographical references and index. | Summary: "This fourth
edition serves as an excellent introduction to general cartographic
principles and as an examination of the best ways to optimize the
visualization and use of spatio-temporal data. It incorporates all the
changes and new developments in the world of maps such as open street
maps and GPS-based crowdsourcing, the use of new web mapping technology,
and adds new case studies and examples. Printed in full color, this
fully-revised edition provides students with the knowledge and skills
needed to read and understand maps and mapping changes, and offers
professional cartographers an updated reference with the latest
developments in cartography"—Provided by publisher.
Identifiers: LCCN 2020015986 (print) | LCCN 2020015987 (ebook) |
ISBN 9781138613959 (hardback) | ISBN 9780429464195 (ebook)
Subjects: LCSH: Cartography.
Classification: LCC GA105.3 .K73 2020 (print) | LCC GA105.3 (ebook) |
DDC 526—dc23
LC record available at https://lccn.loc.gov/2020015986
LC ebook record available at https://lccn.loc.gov/2020015987

ISBN: 978-1-138-61395-9 (hbk)
ISBN: 978-0-429-46419-5 (ebk)

Typeset in ITC Benguiat
by codeMantra

Contents

Preface

PREMISES AND OBJECTIVES

This book has been written to assist in cartographic education and intends, as a first objective, to provide an overview of the role that maps will play both today and in the near future in the world of geospatial data handling. It shows the background against which the provision and visualization of geospatial information takes place. It provides awareness of the Web both as a spatial data source and as a means for distributing the results of visualizing this spatial information. To realize that first objective, the nature of geospatial data is described as well as the characteristics of maps and the ways in which they can be put to use. A development stimulated by the Web was the increased use of spatial data infrastructures, for sharing national and global geodata with the professional and general public. The development of the Internet has boosted the possibilities for interaction and for querying the databases behind the maps presented there. The number of databases available via the Web has increased dramatically and so has the ability to interact with them (query, process, etc.) online. Maps have acquired an important interface function in this new cyberspace geo-information distribution environment. If mapmaking with GIS (geographical information system) mainly involved geo-professionals, the World Wide Web potentially allows everyone to have access to this new medium to create maps.

But not everyone is aware of the intricacies of map design and of the characteristics of the various map types and their limitations. That is where our second objective comes up: teaching map design. What types of geovisualization are appropriate, and how do we translate the numbers collected through censuses or the data measured by sensors into images that allow us to draw sensible conclusions? The answers are shared out over topographic, statistical (or thematic) and temporal maps. For all three categories of maps, we intend to provide sufficient relevant knowledge of cartography and geovisualization concepts and techniques to those accessing the World Wide Web for the production and use of effective visualizations of geospatial information.

Showing the manner in which maps function, either independently or combined in atlases, and can be analysed and interpreted, either in stand-alone or in geo-information environments, is the third main objective of this book. Since the position of the Web has strengthened and stabilized itself, it also stimulated a more integrative approach to problem-solving with geo-information (GIScience (geographical information science)). Since the World Wide Web is highly interactive, and since it allows one to integrate data files, and to link distributed databases, this makes maps suitable instruments for exploring these databases.

BOOK HISTORY AND ACKNOWLEDGEMENTS

Most of the design processes advocated in this book are still based on the inspiration provided by Jacques Bertin's book *Semiology of Graphics* (2011). He wrote the original French version in 1967 in order to improve the printed maps he was confronted with in the media. The media have changed, and the Web is our new medium now, but the basic cartographic design rules still apply in the new interactive visualization of geospatial data for the Internet.

The first, 1996, edition of this book published by Longman was developed from a book (*Cartography: Design, Production and Use of Maps*) for cartography students in the Netherlands, published by Delft University Press in 1993. Since then, a Polish edition has been published (1998) as well as a second edition of the Dutch version (1999). Each of these editions influenced the subsequent editions in other languages. The second English edition was published by Prentice Hall in 2003, and a Russian translation was published in 2005, followed by an Indonesian translation in 2007; an inexpensive English edition for the South-East Asian market is already in its second print. In 2010, Pearson published a third edition, and a year later, The Guilford Press published a larger-sized edition for the United States. A Chinese edition was published in 2014.

The illustrations in these books were initially produced by practical cartographers from Delft and Utrecht universities; the illustrations of the third edition have been based on them but have been reprocessed and updated by Wim Feringa of University of Twente, Enschede, the Netherlands. His was the major job now to convert all illustrations to colour mode, as the publishers allowed us to make this fourth edition a full-colour one, so that the full visual impact of maps could be unleashed.

Though the map examples include many references to situations outside Europe, most of them stem from inside that continent, or even the Netherlands. There are some favourite spots where we return again and again, as many maps refer to Maastricht (the Dutch city where in the year 1992 the treaties were signed that led to the creation of the European Union). Over 20 maps refer to the English Lake District, and another favourite spot is Mount Kilimanjaro. In a sense, these illustrations reflect the professional practice and the related movement patterns of the authors.

STRUCTURE

The basic structure of this book has been left unaltered: it has three distinctive parts. The first five chapters offer the context and basics of maps. The second three each deal with the components of geospatial data: location, attribute and time. The last three chapters deal with 'maps at work' and demonstrate how maps and atlases can assist in problem-solving and decision-making. These three parts are structured as follows: in the first part, in Chapter 1, we discuss the place of maps and mapping in the geo-information environment (GIS (geographical information system), GIScience (geographical information science) and the geospatial data infrastructure, of which the Web constitutes an ever more important part). We proceed to show how data are collected (Chapter 2) and present the concepts that are valid in mapping (Chapter 3) and GIScience (Chapter 4). Chapter 5 deals with the necessary analysis of geospatial data prior to their visualization and also offers some basics of map production.

In the second part, Chapter 6 is focused on location. It not only deals with the characteristics of the base map (reference system, projection, relief portrayal, generalization and geographical names) but also deals with the organizational aspects of topographic map production. Chapter 7 shows the visualization options of the attribute data that are to be rendered on these base maps (thematic map types). Chapter 8 discusses the temporal component of geospatial data.

In the third part, the subject matter becomes more advanced as we deal with the intricacies of map use, analysis and interpretation: Chapter 9 describes how to work with maps and atlases, Chapter 10 shows how to work with maps in a highly interactive geovisualization environment, and finally, Chapter 11 deals with maps for decision-making in a wider context.

UPDATING AND ACCESS

In this age and time where new software packages, new institutional set-ups and technical advances impact us almost continuously, coupled to the more ephemeral aspects of the Web, to keep up to date is a challenge. Although the structure of this book remained the same (the only new sections in this edition are on cartographic education, map machines and story maps), its existing sections were also overhauled. The antithesis between commercial packages that try to monopolize geospatial data handling and public initiatives, such as OpenStreetMap, OpenNauticalChart, Open Geospatial Consortium and the World Wide Web Consortium that try to effectuate the reverse, is dealt with. The launch in 2020 of new GPS platforms and the advent of higher resolution lidar or laser scanning are examples of new techniques described. Usability is a new concern we deal with, as well as volunteered geographical information. Volunteers meet in mapathons, in missing map projects to help out in emergency situations. And the goal we ultimately map for, a sustainable future, where big data are analysed to realize our sustainable development goals is also made explicit. Both authors have been engaged in relevant projects undertaken by the United Nations.

As we have rewritten the text in order to accommodate new generations of web browsers, we also use the prime function of the Web to keep this book up to date. Apart from the Web, every chapter has its section on books for further reading, while all the references to printed literature are grouped together at the back.

From a society that was used to having free access to printed maps, we have evolved to a society used to having free access to geospatial data and maps on the Internet. Everyone can process and visualize the geospatial data available there and put the resulting maps on the Web in turn – there is no quality standard against which the material is checked first, before incorporation is permitted, which is acceptable because the very impact of the Web stems from the fact that it is a free medium. But geo-professionals – and cartographers belonging to this group – have the responsibility of convincing as many as possible to keep the tenets of good and responsible design while visualizing the geospatial data, in order to support the process of spatial decision-making; this refers to a large part of all the cybernetic processes. The decisions based on visualized geospatial data remain only as good as the data and the visualizations themselves.

Menno-Jan Kraak
Ferjan Ormeling
December 2019

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Acknowledgements

We are grateful to the following for permission to reproduce copyright material: Figure 1.6, Rijkswaterstaat NL (CC); Figure 1.10, Hootsmans & Van de Wel, 1993; Figure 1.12, Kadaster Geoinformatie NL (CC); Figure 1.14, Kadaster Geoinformatie NL (CC); Figure 1.15, Open Topo NL (CC); Figure 1.18, Google LLC; Figure 1.19, Google LLC; Figure 2.4, OpenNauticalChart.org (CC); Figure 2.5, Kadaster Geoinformatie NL (CC); Figure 2.14-2.16, Hootsmans & Van de Wel, 1992; Figure 3.6, ANWB Den Haag; Figure 3.13, Bundesamt für Seeschifffahrt und Hydrographie Hamburg/Rostock; Figure 3.14, Kramers, 2020; Figure 3.15, Google LLC; Figure 3.16, Nationale Atlas NL (CC); Figure 3.17, Rijkswaterstaat NL (CC); Figure 3.18, Georg Westermann Verlag; Figure 4.1, Kadaster Geoinformatie NL (CC); Figure 4.8, Koussoulakou 1990; Figure 4.11, The Copyright Licensing Agency Ltd HMSO; Figure 5.23, Patterson; Figure 5.34, Microsoft; Figure 6.15, Noordhoff; Figure 6.17, Open Street Map (CC); Figure 6.24, Kadaster Geoinformatie NL (CC); Figure 6.25, Open Topo NL (CC); Figure 6.31a, EDK; Figure 6.31b, Noordhoff; Figure 6.32, National Geographic; Figure 6.37, Kadaster Geoinformatie NL (CC); Figure 6.39, Ordnance Survey UK; Figure 6.40, Kadaster Geoinformatie NL (CC); Figure 6.41, French Mapping Agency IGN; Figure 6.42, National Land Survey Open data SF (CC); Figure 6.43, USGS; Figure 6.44, Swiss Topo; Figure 6.45, LGL Baden Wurtemberg; Figure 6.46, ESRI; Figure 6.47, Open Street Map (CC); Figure 6.49, Royal Mail (UK); Figure 6.50, Bolder; Figure 6.51, Kadaster Geoinformatie NL (CC); Figure 7.4, Kadaster Geoinformatie NL (CC); Figure 7.16, Freitag 1992; Figure 7.26, Erkamov & Cholodnyj, 1929; Figure 7.36, Dorling, 1995; Figure 8.4, Iceland Air; Figure 9.1, Noordhoff Figure 9.4, Swiss Topo; Figure 9.5-9-6, Microsoft; Figure 9.7, Centennia Historical Atlas; Figure 9.9, Georg Westermann Verlag; Figure 9.10, ANWB Den Haag; Figure 9.13, Centennia Historical Atlas; Figure 10.3, Google LLC; Figure 11.2, ONC US Government; Figure 11.2, ONC US Government; Figure 11.5a, Alphen aan den Rijn municipality; and Figure 11.5b, Falkplan BV.

In some instances, we have been unable to trace the owners of copyright material, and we would appreciate any information that would enable us to do so.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

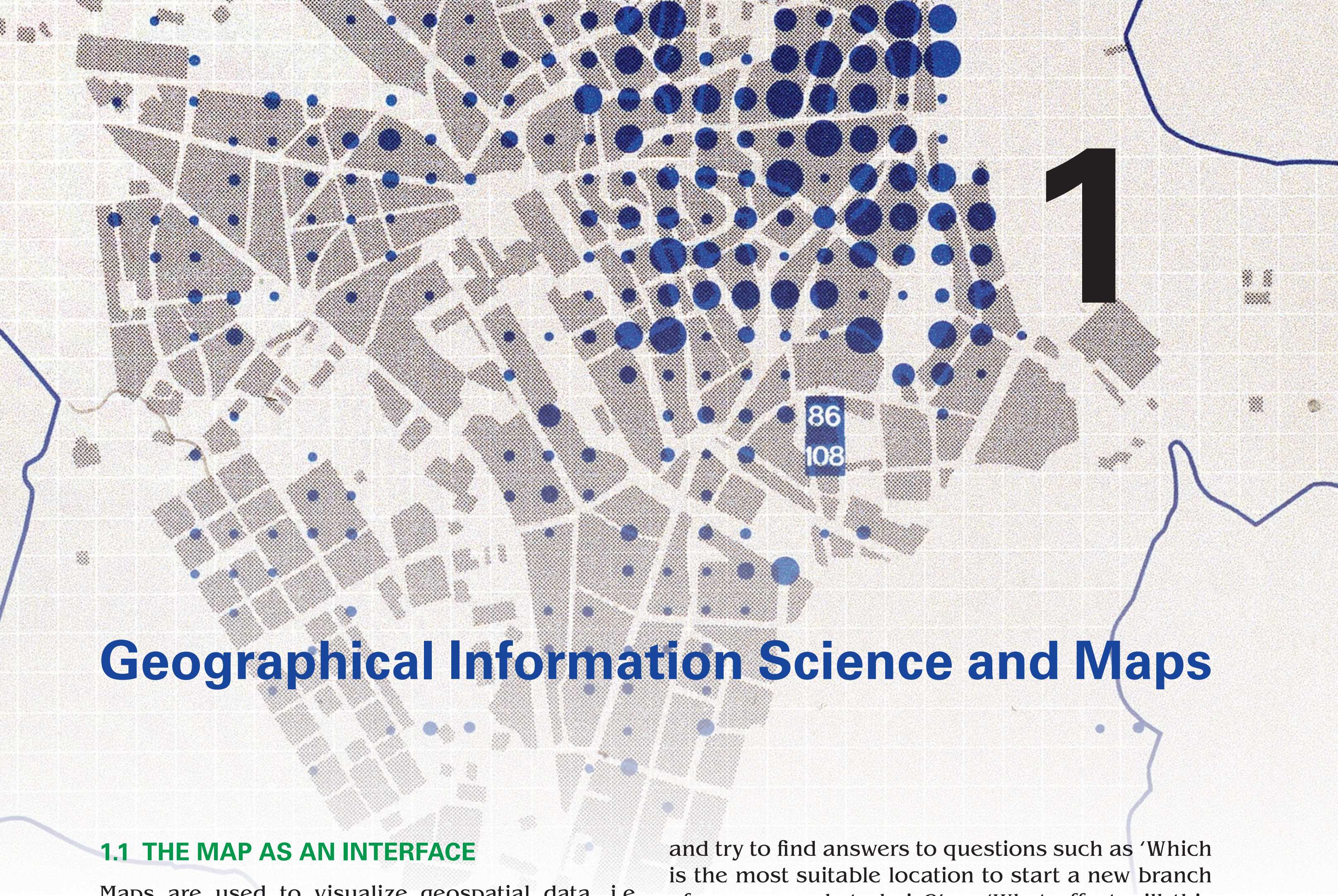
Authors

Menno-Jan Kraak is a professor of geovisual analytics and cartography at the University of Twente/ITC. Currently, he is the head of ITC's Geo-Information Processing Department.

He was the president of the International Cartographic Association (ICA) for the period 2015–2019, and currently immediate past president, and chair of the UN-GGIM Geospatial Societies (2018–2020). He wrote more than 200 publications, among them the book *Mapping Time*. He is a member of the editorial board of several international journals in the field of cartography and GIScience, and the chair of the Foundation Scientific Atlas of the Netherlands.

Ferjan Ormeling worked as an atlas editor before being nominated professor of cartography at Utrecht University. Currently, he is employed at the University of Amsterdam as a member of the Explokart research group. He chaired the Education Commission of the International Cartographic Association (ICA) for 11 years and was the vice-chair of the UN Group of Experts on Geographical Names 2007–2017. In 2017, he was elected president of the UN conference on the Standardization of Geographical Names in New York.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com



1

Geographical Information Science and Maps

1.1 THE MAP AS AN INTERFACE

Maps are used to visualize geospatial data, i.e. data that refer to the location or the attributes of objects or phenomena located on Earth (the terms 'spatial data' and 'geographical data' will be used interchangeably). Maps help their users to better understand geospatial relationships. From maps, information on distances, directions and area sizes can be retrieved; patterns revealed; and relations understood and quantified. Since the 1980s, developments in digital geospatial data handling have gained momentum. Consequently, the environment in which maps are used has changed considerably for most users. With the computer came on-screen maps. Through these maps, the database from which they are generated can be queried, and some basic analytical functionality can now be accessed through menus or legends. In the 1980s, these software packages that allowed for queries and analyses of geospatial data became known as 'geographical information systems' (GISs). As their functionality matured, their application spread to all disciplines working with geospatial data. GIS introduced the integration of geospatial data from different kinds of sources. Its functionality offers the ability to manipulate, analyse and visualize the combined data. Its users can link their application-based models to the data contained in the systems, and try to find answers to questions such as 'Which is the most suitable location to start a new branch of a supermarket chain?' or 'What effect will this plan, or possibly its alternative, have on the surrounding area?'

Maps are no longer only the final products they used to be. The paper map functioned, and functions, as a medium for storage and presentation of geospatial data. The introduction of on-screen maps and their corresponding databases resulted in a split between these functions. To cartographers, it brought the availability of database technology and computer graphics techniques that resulted in new and alternative presentation options such as three-dimensional and animated maps. Geospatial analysis often begins with maps; maps support judging intermediate analysis results and presenting final results. In other words, maps play a major role in the process of geospatial analysis.

The rise of Internet brought the next revolution in mapping. Access to interactive maps is no longer limited to professionals. Products such as Google Maps/Google Earth even allow people to add their own data to the maps and share it with others in a mouse click. The IT-related developments have resulted in a convergence of the different disciplines working with geo-information. GIS is integrated in the workflow of geo-related problem-solving. The disciplines studying related methods and techniques have converged under the header of geographical information science (GIScience). Scientists in this field do research on GIS (e.g. study principles on which GIS is based) and with GIS (e.g. study how GIS can be used in scientific applications (Longley et al., 2015).

The above development also led to spatial or geographical data infrastructure (SDI or GDI). Next to a technical setting, a GDI comprises a set of agreements and arrangements to access, integrate and use geo-information. These new infrastructures for accessing geospatial data are being developed all over the world in order to allow access to the geospatial data files created and maintained in order to monitor the population, resources and environment spatial aspects of our modern societies. Access to the data needed requires complex querying procedures that are simplified when using maps to pinpoint the areas and themes for which data are needed (Figure 1.1).

In a GIScience environment, visualization is applied in four different situations. Firstly, visualization can be used to explore, for instance, in order to play with unknown data. In several applications, like those dealing with remote sensing data, there are abundant (temporal) data available. Questions such as 'What is the nature of the data set?' or 'Which of those data sets reveals patterns related to the current problem studied?' have to be answered before the data can actually be used in a geospatial analysis operation. Secondly, visualization is applied in analysis, for instance, in order to manipulate known data. In a planning environment, the nature of two separate data sets can be fully understood (e.g. the groundwater level and the possible location of a new road), but their relationship cannot. A geospatial analysis operation, like overlay, can combine several data sets regarding the same area to determine their possible geospatial relationship. The result of the overlay operation could, when necessary, be used to adapt the plans. Thirdly, maps are used to synthesize the results of the analysis. Fourthly, visualization is applied to present or communicate the new geospatial knowledge. The results of geospatial analysis operations can be displayed in well-designed maps easily understood by a wide audience. The cartographic discipline offers design rules to do so. As the fourth objective of visualization, we have already mentioned the easier access to the data files behind the maps.

Considering these four different fields of visualization in GIScience (exploration, analysis, synthesis and presentation), it can be noticed that the tools for presentation are the most highly developed. While producing maps to communicate

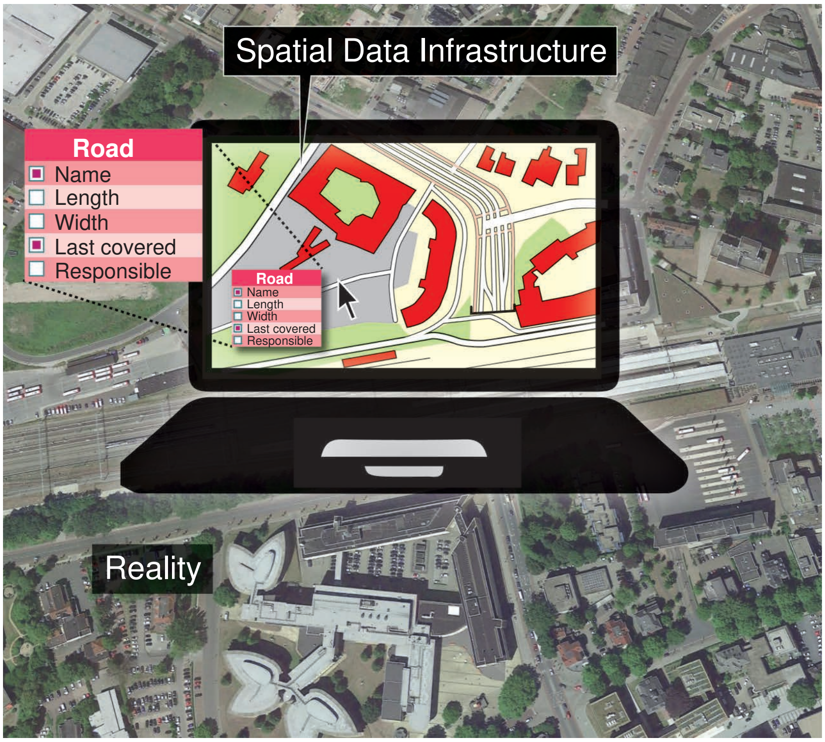


FIGURE 1.1 The interface role of maps in the spatial data infrastructure. Here, the GIS answers a query regarding a clicked object

geospatial information, cartographic rules and guidelines (together called 'cartographic grammar', based on the nature of the data and the communication objectives) are available to make the maps effective. However, as these rules are not part of the mapping software, it allows users to produce their own maps even when they are unaware of cartographic grammar. In other words, there is no guarantee that the maps will be effective. These cartographic rules could also be applied in the analysis phase, but the necessity to do so would be less strong here. When cartographers and analysts discuss this matter, the second group would always claim 'Who cares about mapping rules, as long as one understands one's own maps?' And because the analysts knew their own data, they probably would understand their own maps, but when showing their maps to others communication problems would start. In a data exploration environment, it is likely that the user does not know the exact nature of the data and therefore might not be able to apply the relevant cartographic rules.

At this moment, the terms 'private visual thinking' and 'public visual communication' should be introduced (DiBiase, 1990). Private visual thinking refers to the situation where users explore and analyse their own data, and public visual communication refers to the situation where users present their results in the form of maps to a wider audience. The first describes the exploration circumstances, and the second presentation circumstances. Analyses can be found somewhere in the middle along a line between the two. This becomes more evident when it is realized that private versus public map use (i.e. maps tailored to an individual versus those designed for a wide audience) is just one of the axes of the so-called map-use cube, first introduced by MacEachren (1994). Along the two other axes, the revelation of the unknown versus the presentation of the known, respectively, and high versus low user interaction are plotted, which are shown in Figure 1.2.

Most chapters in this book concentrate on maps that should communicate geospatial information (the lower-left front corner of the cube). However, recent developments in cartography and other disciplines handling geospatial data not only require a new line of thought, but also create one. This can be illustrated by plotting the evolutionary stages of the development of electronic atlases in the cube along the diagonal from the corner 'wide audience, presenting knowns and low interaction' towards the corner 'private use, presenting unknowns and high interaction' (Figure 1.2b). Possibilities for interaction are boosted by the advent of the Internet and its potential for querying the databases behind the maps presented there. Early electronic atlases were, in effect, sequential slide shows, but today's electronic atlases have high interactive multimedia mapping capabilities, and allow users to combine

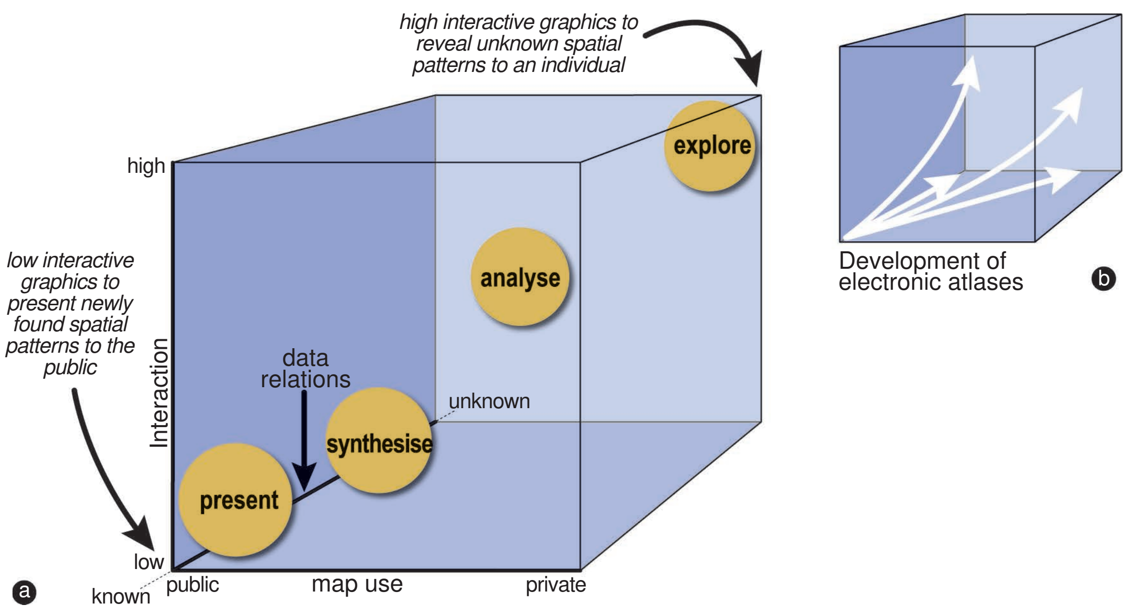


FIGURE 1.2 The map-use cube (adapted from MacEachren and Taylor, 1994): (a) the four main situations to visualize data in a GIS, (1) to present, (2) to synthesize, (3) to analyse and (4) to explore; (b) the evolution of the electronic atlas since 1987 plotted in the map-use cube

their own data with atlas data. Each category of map use in Figure 1.2 cube asks for its own visualization approach. New cartographic tools and rules have to be found for these approaches. They are probably not as restrictive as traditional cartographic rules, but on the other hand not as free as the technology allows either.

The demand for sophisticated geospatial data presentation is further stimulated by developments in scientific and information visualization, multimedia, virtual reality and exploratory data analysis. In each of these external developments influencing GIScience and maps, it would appear that from a technical point of view, there are almost no barriers left. The user is confronted with a screen with multiple windows displaying text, maps and even video images supported by sound. Important questions remain. Can we manage all the information that reaches us? The ever more detailed satellite imagery available, the increasing number of sensor networks and new techniques for analysing textual sources with spatial references like geoparsing all lead to highly varied 'big data', characterized by large volumes of data, coming available with high velocity. In order to make sense of them and derive meaning or trends, cartography, with its capacity of generalizing data in order to fit their purpose, plays an important role. What will be the impact of these developments on the map in its function to explore, analyse and present geospatial data? This book tries to provide an overview of the role that maps will play both today and in the near future in the world of geospatial data handling. There is an enormous amount of geospatial data out there, on the Internet, useful for any kind of geospatial research, waiting to be harnessed, made accessible and structured by being visualized as maps. The nature of these geospatial data is discussed in Chapter 2, and the characteristics of the maps that visualize them are dealt with in Chapter 3.

1.2 GEOSPATIAL DATA

Geographical information is different from other information in that the data, as a special characteristic, refer to objects or phenomena with a specific location in space and therefore have a spatial address. Because of this special characteristic, the locations of the objects or phenomena can be visualized, and these visualizations – called 'maps' – are the key to their further study. Figure 1.4 shows how objects from the real world that can be localized in space (such as houses, roads, fields or mountains) can be abstracted from the real world as a digital landscape model (DLM), according to some predetermined criteria, and stored in GISs (such as points, lines, areas or volumes) and later, after being converted into a digital cartographic model (DCM), represented on maps (with dots, dashes and patches) and integrated in people's ideas about space. When stored in a database, these geospatial data are usually divided into locational data, attribute data and temporal data. The first refers to the geometrical aspects (position and dimensions) of the phenomenon one has (geometric) information about, and the second refers to other, non-geometrical characteristics. The temporal data refers to the moment in time for which both the locational and attribute data are valid. These three aspects are linked to the elementary questions 'Where?', 'What?' and 'When?' (see Figure 1.3a), and define the nature of an object (Figure 1.3b). An object's location, attribute or time can have multiple characteristics, such as different coordinate systems, multiple variables and even different kinds of time (Figure 1.3c). Apart from these three questions, one might also ask 'Why?' or 'How?' Answering these last two questions requires further analysis of the data. It might require more attention from one of the data's components, resulting in a perspective from what could be called 'location space', 'attribute space' or 'time space' (Figure 1.3a). Chapters 6–8 will, respectively, deal with questions related to these three spaces.

The stored geospatial data on a specific study area is called its 'digital landscape model' (DLM). Of course, it is an abstraction: selected characteristics have been measured or assessed and incorporated in this DLM. As soon as this DLM is considered suitable for communication to other persons, and has to be produced in hard copy form, this model has to be converted into a DCM, which consists of a series of instructions to the printer or screen, to produce dots, dashes or patches, in different sizes, colours, etc. for multiplication and distribution (Figure 1.4). Finally, users of the mapped information will view it and process it into their cognitive map, the mental construct of space they will base their decisions on.

For data to qualify for the tag 'geometric data 'or 'georeferenced data', information about their location would be required. This can be geographical or reference grid coordinates, code numbers that refer to statistical areas, topological terms (e.g. A is in between B and C) or nominal terms, as in street addresses and postcodes. The geospatial nature of the objects can be expressed in their shapes, with which one represents objects from the real world. There is a basic subdivision into point-, line-, area- or volumetrically shaped objects

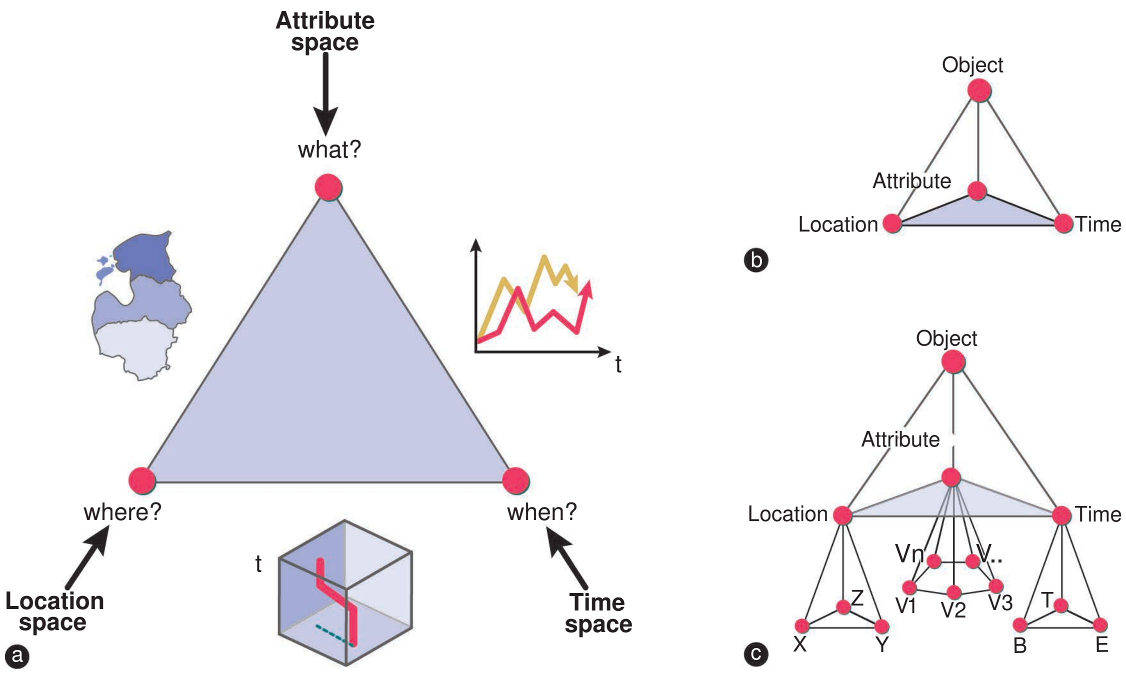


FIGURE 1.3 The characteristics of geospatial data: (a) its components' location, attribute and time, and their related elementary questions 'Where?', 'What?' and 'When?'; (b) the object view; (c) detailed characteristics of the data components

(see Figure 1.5), and this can be further subdivided into, for instance, elongated, triangular, irregular or convex-shaped objects. In a sense, this is scale- or resolution-dependent, as a populated settlement will be rendered by a point in a national context and as a built-up area in a municipal context.

Whether the objects or phenomena from the real world are abstracted as discrete or continuous is very important for subsequent storage and mapping procedures. Discrete objects can be bordered on all sides, and the coordinates of these borderlines can be made explicit. These can either be the locations of tactile objects (houses, streams, etc.) or be the locations of predetermined areas (states, enumeration areas or distribution areas). Continuous representations are abstractions of those phenomena that are considered to change non-incrementally in value. They can be tactile or measurable (like precipitation data or gravity field data) or be based on models (like isochrones, i.e. lines linking points that can be reached in an equal travelling time from a given starting point).

For later visualization procedures, it is essential that the nature of the attribute information be established. These attributes can refer to visible characteristics (e.g. deciduous trees) and invisible characteristics (e.g. temperature). When attempting to define these attribute values of objects, one usually tries to measure or categorize them, and then, it will appear that these characteristics are either qualitative or quantitative. One may distinguish a number of measurement scales (see also Chapter 5), on which the values for these characteristics can be assessed:

- *Nominal scale*: Attribute values are different in nature, without one aspect being more important than another (e.g. different languages or different geological formations).
- *Ordinal scale*: Attribute values are different from each other, but there is one single way to order them, as some are more important/intense than others (e.g. warm, mild, cold).
- *Interval scale*: Attribute values are different and can be ordered, and the distance between individual measurements can be determined. Temperature is a good example: because the respective zero points of their measurement scales have been selected at random, it is impossible to say that, for instance, a temperature of 64°F is twice 32°F. This is plain when the values are converted into Celsius and become 18°C and 0°C, respectively.
- *Ratio scale*: Attribute values are different and can be ordered. Distances between individual measurements can be determined, and these individual measurements can be related to each other. If, for instance, the gross domestic

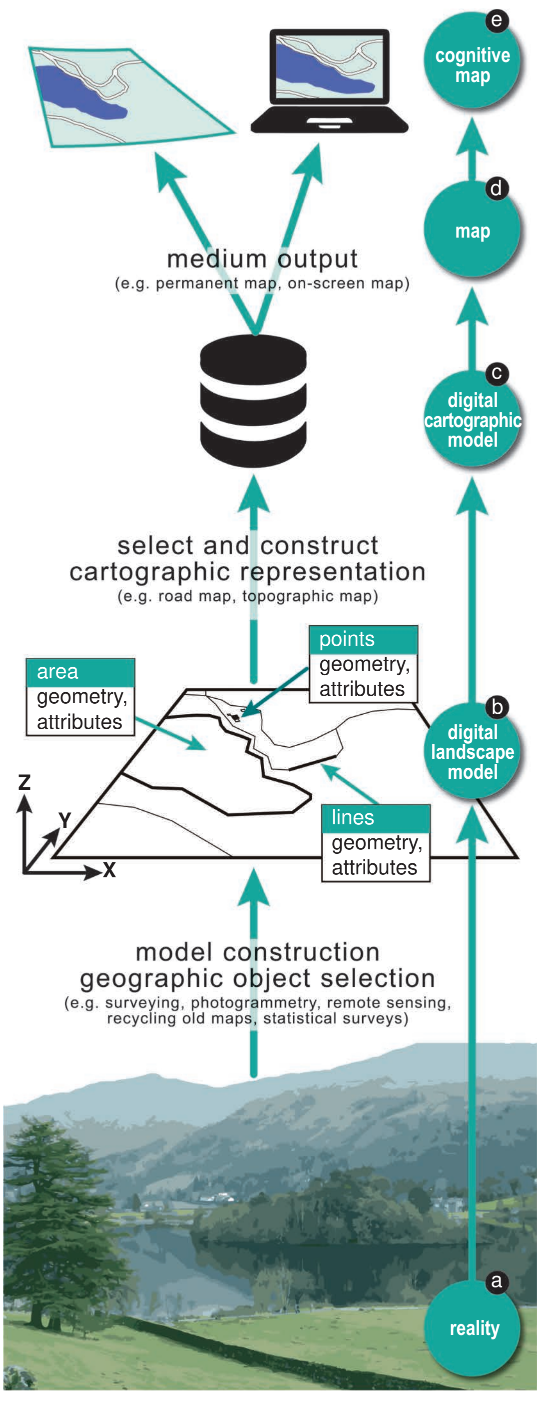


FIGURE 1.4 The nature of geospatial data: from reality (a), via model construction and selection to a digital landscape model (b), followed by selection and construction of a cartographic representation towards a digital cartographic model (c), presented as a map (d), which results in the user's cognitive map (e)

income per capita in Sri Lanka in 2017 was $13 000 per annum and in Bangladesh $4300, then one can say that the amount in the former was thrice the value of that in the latter.

All the geospatial data will be subject to changes over time: the attribute information on an object can change over time (like the composition of the population of an area), and even the object's location itself may change (for instance, through continental drift). The data's timestamp is seen as the third major component, next to geometry and attribute values. Especially these days, the interest in the data's temporal component increases because of the expanded number of time series available and the wish to analyse processes over time instead of during a single time slice.

One is only able to study or analyse or interpret geospatial data after the data have been visualized, i.e. after the application of the cartographic grammar to render these objects and their relationships. Here, symbols and signs are used, i.e. dots, dashes and patches, and these can vary in size, shape, texture, colour, value and orientation (see Chapter 5). These signs are linked to the objects or relationships they represent, and by doing so, one is able to convey geospatial relationships between point, line, area or volumetric objects, in a number of dimensions, to the map user.

If only one dimension is available, then the location of geospatial data can be expressed, for instance, by their distance from a central market, or from a point of origin – represented as a straight line. Two-dimensional representation with these dots, dashes and patches will result in a planimetric map, and Figure 1.6a showing a contour map of a hill in the south of the Netherlands is a good example. In order to have a true three-dimensional representation, a physical model of this hill could be produced from cardboard, or a virtual model could be created, which, by rotating it, or through anaglyphs, using red and green glasses, could be seen from all sides on a monitor screen. By drawing the model in perspective, and using a hidden-line algorithm, this 3D aspect could be simulated (represented in Figure 1.6b). The current description of this type of rendering is '2.5 dimensional', because it is a projection of 3D reality on a 2D plane but still gives the map user a three-dimensional impression. Maps with hill shading (see Figures 6.30 and 6.31) are another example of this.

If one adds the time dimension, the representation will become four-dimensional (Figure 1.6c), when, through juxtaposition of two states of this object, for instance, in 1950 and 2020 respectively, the change in its geometry or attributes during the intervening period can be ascertained.

Through their visualization in maps, the geospatial relationships of objects can be made visible. These geospatial relationships will usually refer to

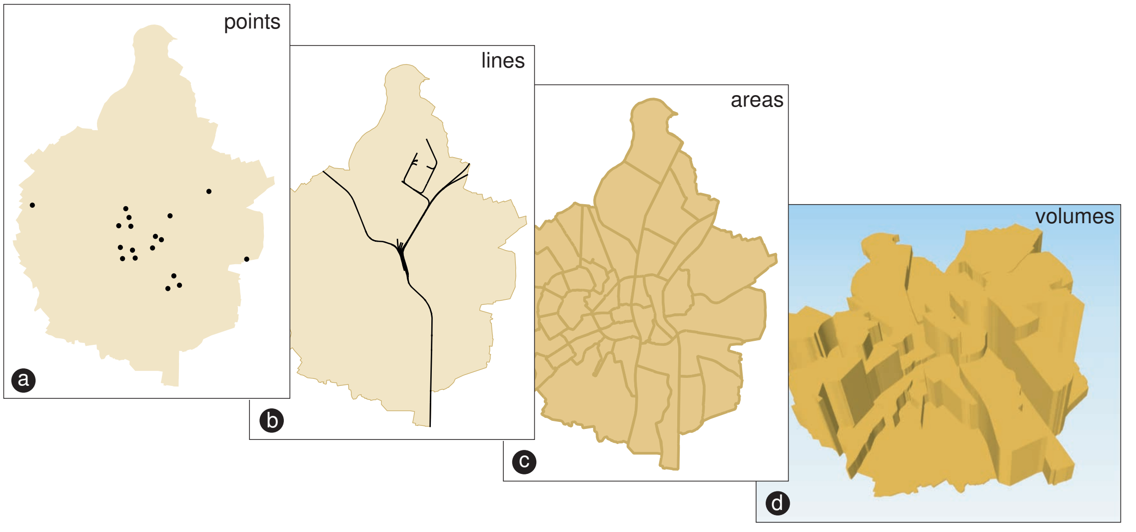


FIGURE 1.5 The representation of geographical objects in a digital environment as (a) points, (b) lines, (c) areas and (d) volumes

relations to some specific location on the Earth's surface, or can be those of objects to one another, and these relations can have many forms. Primary geospatial relationships are those between objects and their location on Earth, or between these objects and their attributes, such as the type of vegetation occurring at that location and the type of road classification. By visualizing object categories from a data file (e.g. car factory locations or stream networks, or fields used for horticulture), relations between the elements for an object category will be made clear, and one will be able to perceive patterns or geospatial trends. By combining geometric and attribute data, one would be able to perceive how the locations of elements from different object categories might influence each other.

Next to these primary relationships, it is possible to perceive secondary types: relations of objects to linear or area reference units, such as that of inhabitants to surface area, the number of cars to the length of the highway network and the relative amount of horticulture in all agricultural areas. One could go further and introduce other dimensions (like height or time), so that tertiary or even higher-order relationships would emerge.

1.3 GEOGRAPHICAL INFORMATION SYSTEMS

For most disciplines working with geospatial data, one of the first uses of the computer was to create an inventory of discipline-dependent data. In this period, cartographers worked to build a database from which they could produce the maps that were previously created manually. In the next phase, spatial analysis of the collected data was emphasized. Forestry scientists, for instance, would apply statistical methods to the maps' attribute data. For cartographers, this meant the possibility of creating different derived products from the existing database. Nowadays problems are approached in an interdisciplinary way. In physical planning processes or in environmental impact studies, data from many different fields are needed. The need to combine them led to the development of GIS. Cartographic knowledge is used in GIS to create proper visualizations. GIS offers the possibility of integration of geospatial data sets from different kinds of sources, such as surveys, remote sensing, statistical databases and recycled paper maps. Its functionality allows one to manipulate these data, or to set up geospatial analysis operations in conjunction with application-based models, and they allow for the visualization of the data at any time during this process. The core functionality of a GIS is provided by the disciplines such as geography, geodesy and cartography that are used to work with spatial data. To this core, functionality from database technology and computer graphics is added. Currently, GIS is used in virtually all disciplines and professions that require geospatial data to execute their tasks or solve their problems.

Why is GIS unique? Because it is able to combine geospatial and non-geospatial data from different data sources in a geospatial analysis operation in order to answer all kinds of questions. The fact that those questions can be answered is quite remarkable if one realizes that geographical phenomena

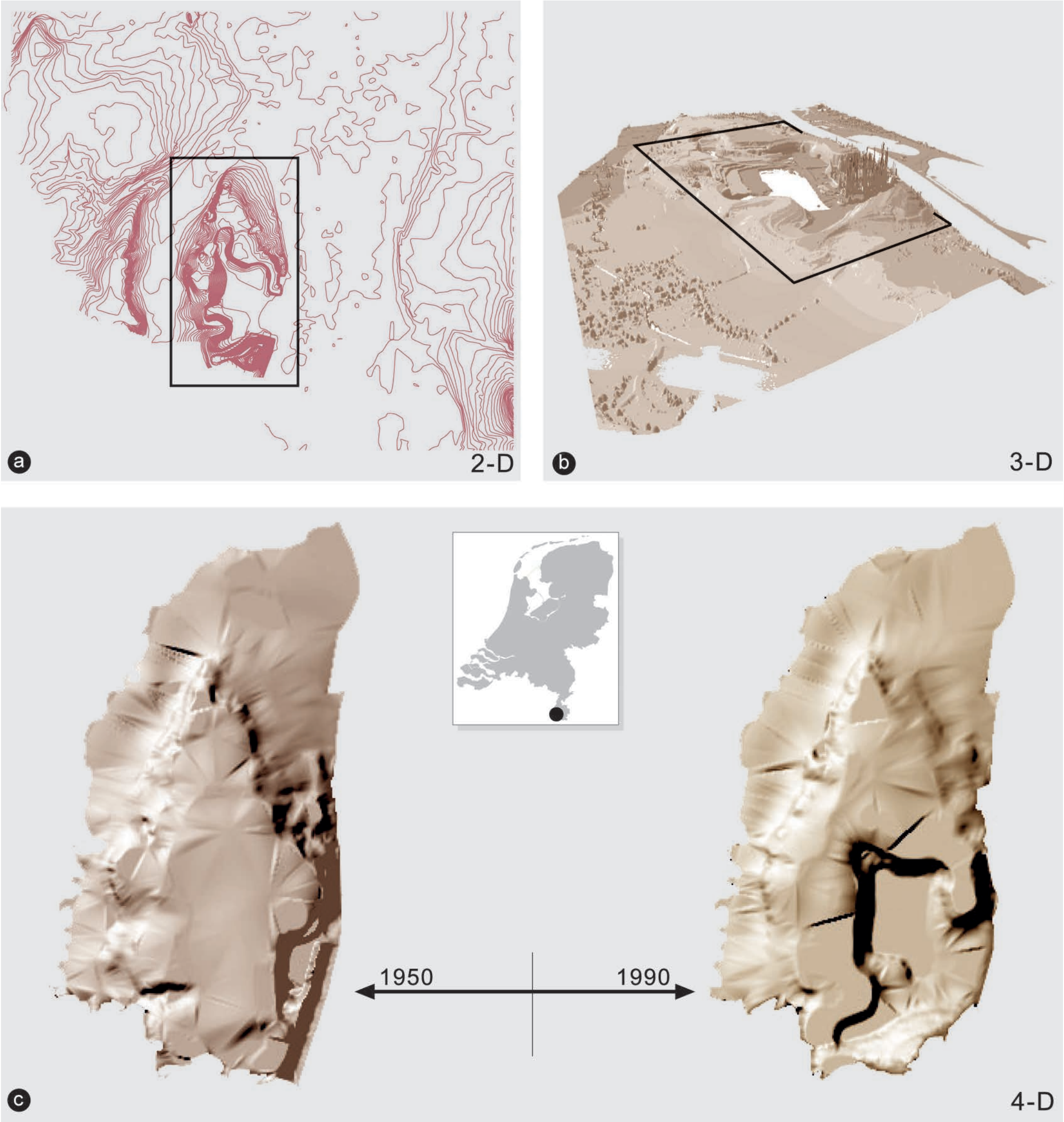


FIGURE 1.6 The dimensionality of geographical objects: (a) (2D), (b) 3D; (c) 4D/time. In Figure 1.6c, part of the hill has been excavated since 1950 (Rijkswaterstaat AHN (CC0))

are almost never homogeneously distributed. And added to this observation, one can add a quote from Tobler (1970) who said: 'all things are related, but nearby things are more related than distant things'. What type of questions can be answered by a GIS (Figure 1.7)?

- *What is there ...? Identification*: By pointing at a location on a map, a name, or any other attribute information stored on the object, is returned. This could also be done without maps, by providing the coordinates, but this would be far less effective and efficient.
- *Where is ...? Location*: This question results in one or more locations that adhere to the criteria of the question's conditions. This could be a set of coordinates or a map that shows the location of a specific object, or all buildings in use by a certain company.
- *What has changed since ...? Trends*: This question includes geospatial data's temporal component. A question related to urban growth could result in a map showing those neighbourhoods built between 1950 and 2020.
- *What is the best route between ...? Optimal path*: Based on a network of paths (e.g. roads or a sewage system), answers to such queries for the shortest or cheapest route are provided.

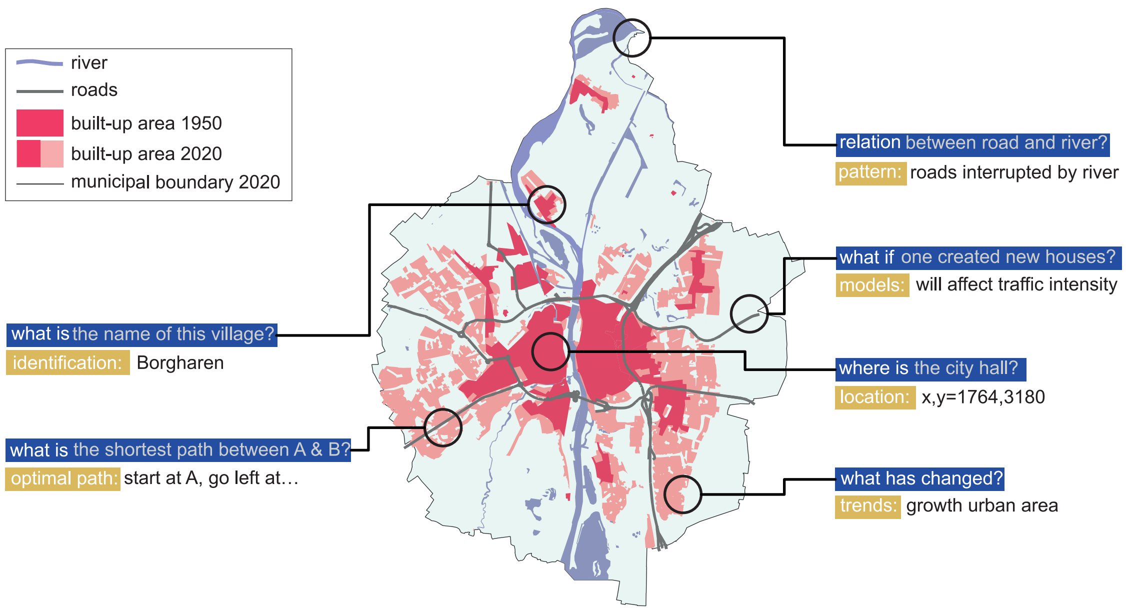


FIGURE 1.7 Typical GIS questions answered by maps such as those used to identify, locate or find geospatial patterns

- *What relation exists between …? Patterns*: Questions like this are more complex and often involve several data sets. Answers could, for instance, reveal the relationship between the local microclimate and location of factories and the social structure of surrounding neighbourhoods.
- *What if …? Models*: These questions are related to planning and forecasting activities. An example is: 'What will be the need to adapt the local public transport network and its capacity when a new neighbourhood is built north of the town?'

Of course, one does not only query a GIS but also uses it interactively, for instance, in physical planning procedure, through manipulation of designs, etc.

GIS development was stimulated by individual fields such as forestry, defence, cadastre, utilities and regional planning (see Chapter 4). Since they all have different backgrounds and different needs, the functionality of the software GIS initially used was different as well. It ranged from statistical analysis packages to computer-aided design packages. Functionality was added, and each of these groups started to call their software a 'GIS package'. This resulted in different meanings for the same term. Next to GIS, literature offers wordings such as land information systems, geo-based information systems, natural resources information systems and geodata systems (Longley et al., 2015).

The multidisciplinary background of GIS led to a multitude of definitions. In general, they can be split into two groups: those with a technological perspective and those with an institutional/organizational perspective. An example of the first is the definition by Burrough and MacDonell (1998): 'a powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world'. An example of the second is that of Cowen (1988), 'a decision support system involving the interaction of geospatially referenced data in a problem-solving environment'. So, it is the potential combination of different data sets that is paramount. A working definition for this book is derived from a combination of the two above: a GIS is a computer-assisted information system to collect, store, manipulate and display spatial data within the context of an organization, with the purpose of functioning as a decision support system.

In order to manipulate geospatial data, to procure added value, a GIS consists of software, hardware, geospatial data and people (the organization). These components communicate via a set of procedures. In Figure 1.8, which summarizes the view of GIS adopted in this book, the central schemes present the GIS components: the problem-solving production line (from exploration to presentation), the potential for geospatial analysis (with application-based models) and integrating geospatial data sets, the unique and basic ingredients of the system. The outer shell renders the organization in which GIS functions. The configuration of the

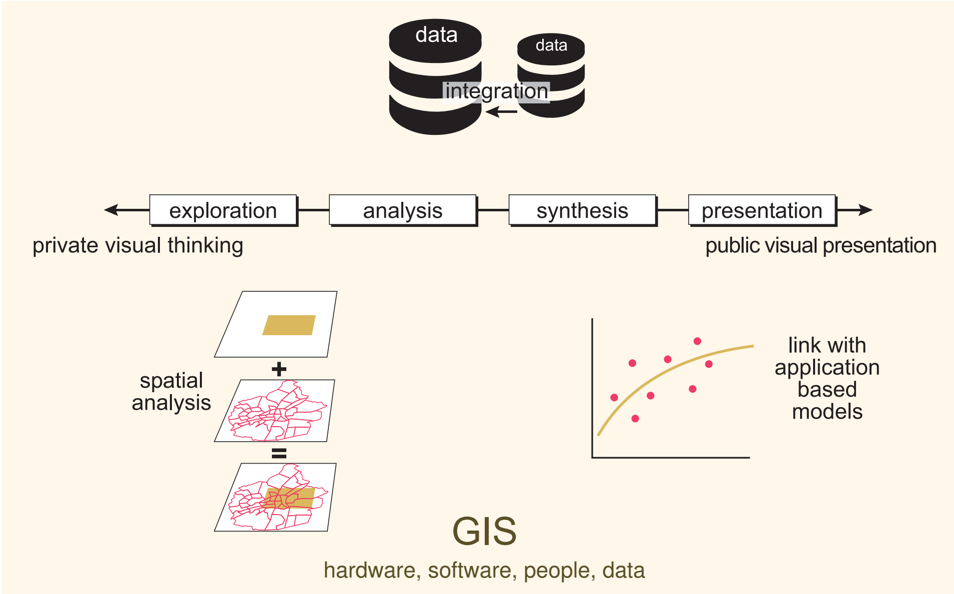


FIGURE 1.8 View on GIS: its characteristics and relation to visualization

scheme stresses the need for a proper user interface and management of the system.

Each organization involved in geospatial data will require a GIS with emphasis on a specific set of functions, depending on its task. In general, functions are needed for data input and encoding (e.g. digitizing, data validation and structuring options), data manipulation (e.g. data structure and geometric conversions, generalization and classification options), data retrieval (e.g. selection, spatial and statistical analysis options), data presentation (mainly graphical display options) and integrated data management.

For the purpose of this book, also written for GIS analysts who have to learn to use the methods, cartography will be regarded as an essential support for nearly all aspects of handling geographical information for the following reasons:

- Maps are a direct and interactive interface to GISs, a sort of graphical user interface with a geospatial dimension;
- Maps can be used as visual indexes to phenomena or objects that are contained in the information systems;
- Maps, as forms of visualization, can help in both the visual exploration of data sets (also the discovery of patterns and correlations) and the visual communication of the results of the data set exploration in GISs;
- In the output phase, the interactive design software of desktop cartography is superior to the output functions of current GISs.

These should be enough reasons for cartography to have an important place in GIS, but there are more reasons, if one looks at the context in which GISs are being used: they are aimed at decision support, and as this regards decisions about geographical objects, it should be visual decision support, in order to take into account the geospatial dimension as well. In order to correctly use these visual decision support aids (the maps visualized on the computer screen or the hard copy output of these systems), the users should adhere to proper map-use strategies (see Chapter 11). This ability to work with maps and to correctly analyse and interpret them is one very important aspect of GIS use. Strangely enough, not one GIS manual gives any clarification in this field, assuming that all GIS users are aware of the ins and outs of map use.

But there is another important decision support aspect to the information that is processed in and presented by a GIS: data quality. GISs are very good in combining data sets; notwithstanding the fact that these data sets might refer to different survey dates and different degrees of geospatial resolution, or might even be conceptually unfit for combination, the software combines them and presents the results. Cartographers, in compiling maps, have worked with different data sets for centuries and have some experience in the transformations that are necessary in order to combine data sets with different resolutions, projections, reference systems, geoids and dates of survey. They have developed transformations and modelling procedures (like generalization) that take account of these differences and will allow for real data integration. They have developed documentation techniques that will describe all relevant data characteristics (metadata) necessary for proper integration to take place. They have also lobbied for decades to standardize these documentation techniques so that the data sets can be easily exchanged (Figure 1.9). So much of the methodology for the determination of data quality is potentially available for GIS users from cartography. Finally, cartographers have strived to have access to geospatial data improved, as is the object of improving the SDI (see Section 1.5).

As to the assessment of data quality, this can be defined, as was shown in the preceding section, as a measure of the suitability of data for specific applications. So, one can, for instance, determine the precision x to which objects (like parcels) in a data set have been localized, as well as the probability p with which these objects have been correctly classified or categorized (e.g. regarding their land cover). Now the GIS will allow the combination of this land cover information (e.g. surveyed in 2018) with precipitation information (e.g. surveyed for the period 1990–2010, for five points in the area, and interpolated for all other locations), with a planimetric accuracy y of the rain gauge locations and a representativity factor z_{1-5} of these measurements for their surrounding areas. Well, what will be the value of correlations between precipitation data and land cover data, taking into account these accuracy, probability and representativity values?

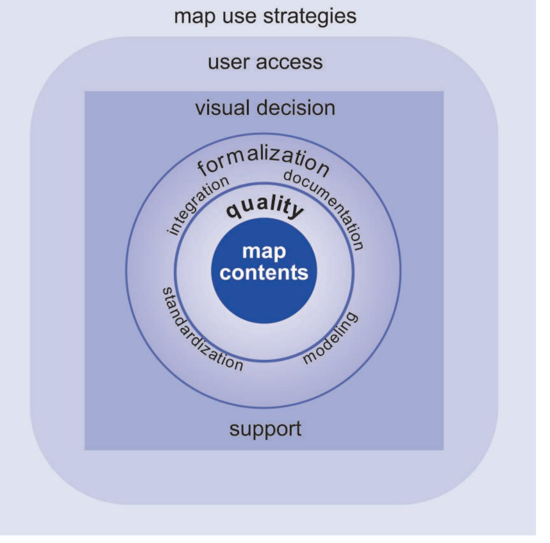


FIGURE 1.9 Visual decision support for spatiotemporal data handling. Keywords in the GIS cartography approach are map-use strategies (how people make decisions based on maps), public access (whether and how people can work with the information), visual decision support (showing what the quality of the information is) and formalization (building expert systems)

Until recently these have been disregarded in GISs, but in a mature GIS that really functions as a decision support system, these values should be indicated to properly inform the decision-maker.

Figure 1.10 shows the classification of remote sensing imagery of urban areas in a part of the

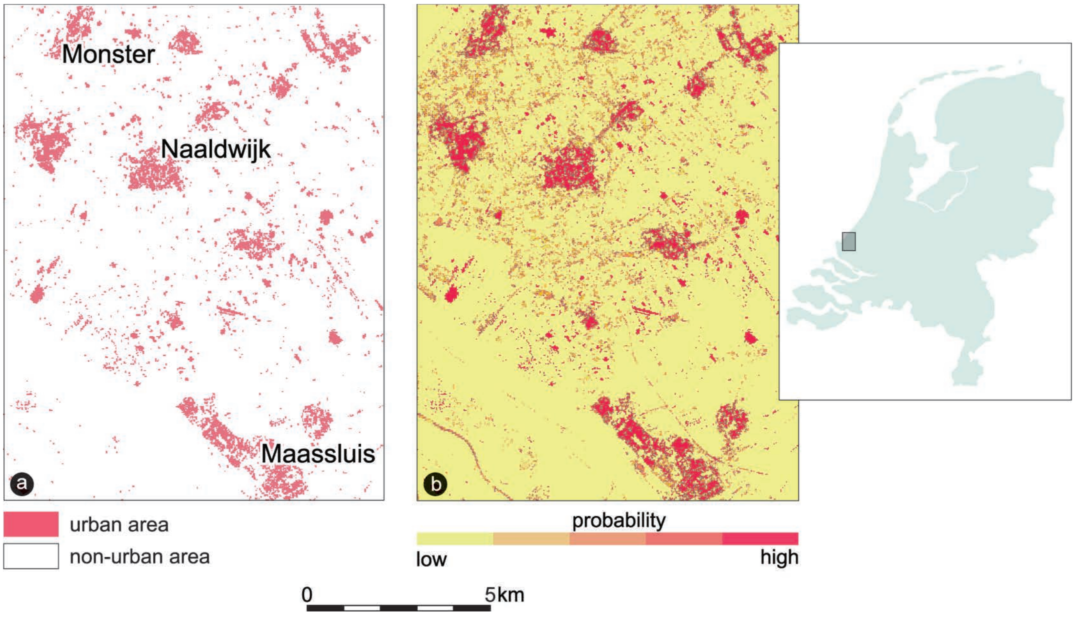


FIGURE 1.10 Geospatial information and meta-information: (a) distribution of urban areas over a part of the western Netherlands; (b) probability map for the classification 'urban' (from Hootsmans and Van der Wel, 1993)

western Netherlands (near The Hague): the image on the left is a classification based on spectral qualities, whereas the image on the right is a probability map for the classification 'urban'. It takes into account the potential confusion with related spectral signatures. The probability of a pixel being correctly classified as 'urban', and not as 'hothouse' or industrial complex or beach or bare soil, can be computed, and visualized probability values for correct classifications like these should form an essential element of the decision-making process, which takes place everywhere where geospatial information is involved.

So, GIS users can be provided with essential tools in all phases of collecting, processing and analysing geospatial data, and communicating it to decision-makers. Those GIS users who able to use maps are provided with the infrastructure for a correct decision-making procedure, and with the necessary information (meta-information) on the quality of the data contained in those maps.

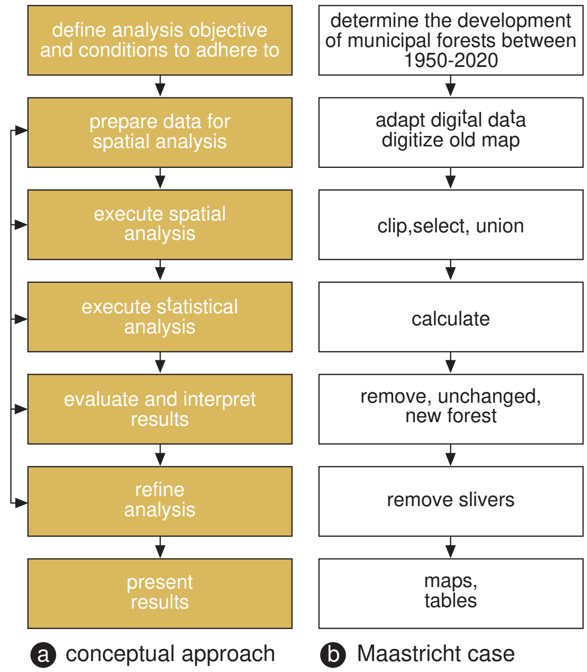


FIGURE 1.11 Spatial analysis: (a) conceptual approach; (b) Maastricht case

1.4 GEOSPATIAL ANALYSIS OPERATIONS

Geospatial analysis operations are the unique processes GIS has to offer to the geospatial data handling community. However, its principles have been known since the 1950s from fields such as quantitative geography and statistical geography (Hägerstrand, 1967). This section will explain the principles of these operations, illustrated with a simple example that demonstrates the strength of GIS and the prominent role maps can play (Figure 1.11).

The first example deals with an issue in the Netherlands municipality of Maastricht. The first step in a geospatial analysis operation should be the definition of its objective and the conditions to adhere to. These conditions can include specific restrictions and constraints. In the Maastricht case, the municipal authorities wanted to know how the municipal forests had developed between 1950 and 2020. They wanted to have a map indicating obliterated, unchanged and new forests, as well as a table with the size of the forest parcels. One of the reasons for this analysis was to check whether a large private company in the municipality had adhered to the conditions agreed upon. The company had a concession to excavate marl in a quarry in the south of the municipality, but it had to plant new trees in those areas that had been affected by its operations.

The second step in an analysis is to prepare the geospatial data. Usually, not all data are available in a format that fits the requirements of the geospatial analysis. Some data have still to be collected in the field, whereas other data must be brought from external sources or brought in from other municipal departments, while there is a fair chance that some of the data are available on paper maps or in tables only. When available digitally, they could still be in a different coordinate system. Other problems that are likely to occur are the data being available on a different aggregation level and the density of coordinates in one set being too low to be compared with another data set; an example of the former problem is statistics being available at a neighbourhood level instead of at a street block level. The main GIS task at this stage is to make sure all data can be integrated and formatted so that they will be fit for use.

To answer its query, Maastricht municipality needed data on land use from both 1950 and 2020 to extract information on the forest parcels. The 2020 data were obtained from a database that contained data from the topographic map 1:25 000. The 2020 municipal boundary was taken from the large-scale map database available. However, it was too detailed to be used in conjunction with the 2020 topographic map data. It had to be generalized first. The 1950 data were not available in a digital form. The original topographic map had to be ordered from the municipal archives and had to be converted into the digital form. An interesting problem that occurred during the conversion was that part of the area for which the data were needed, within the 2020 municipal boundary, was not within the municipal boundaries on the 1950 map, because the area of Maastricht municipality had been enlarged considerably since, as can be seen in Figure 1.12a. The generalized 2020 boundary was used as a digitizing mask to extract the correct information from the 1950 map (see Figure 1.12b). The same 1990 boundary was used to select the land use data from the 2020 topographic map database. This database contained, in several layers, all the data of a topographic map sheet that covers a large area in the southern part of the Netherlands. From it, only data related to Maastricht were needed. Working with the whole database would slow down the execution of the operation because of the large amount of data.

The third step in the operation was the actual execution of the geospatial analysis. Most packages have all kinds of operations available, which somehow operate on the geometric and non-geometric components of the data. In general, one distinguishes three major types of geospatial analysis operations: overlay and buffer operations, network operations and surface operations. The first category often combines several data sets based on certain criteria, the second category uses an infrastructure network to find optimal paths, and the third category determines all kinds of (terrain) surface characteristics. Most packages allow the combination of these operations. The Maastricht case was limited to a simple overlay operation. The data sets from 1950 to 2020 were combined, which revealed those forest parcels unchanged, those removed and the new ones. Figure 1.13 demonstrates this process. From the 1950 to 2020 land use data sets, the forests were selected to create two new data sets, forests 1950 and 2020. It was those two sets that were combined in the overlay operation. This calculated all possible intersections between the 1950 and 2020 forest parcels. In this process, all attributes were inherited and saved in the final data set. Based on these attributes, the map in Figure 1.13e could be drawn. The command 'draw all forest parcels with an attribute year 1950 and not year 2020' would result in all the parcels obliterated since 1950 being drawn.

The statistical analysis of the results is the step after the geospatial analysis. It is executed to fulfil the conditions set when the objective was determined. When, in the Maastricht case, one criterion would have been that the municipality is only interested in those parcels over a certain area or perimeter or when, in addition, it would like to know the average size of the new parcels, some basic statistics could be applied here. In complex geospatial analysis operations, it would be likely that more sophisticated statistical methods are needed. The next step in the analysis was the evaluation and interpretation of the results. In general, when executing a geospatial analysis, one has certain

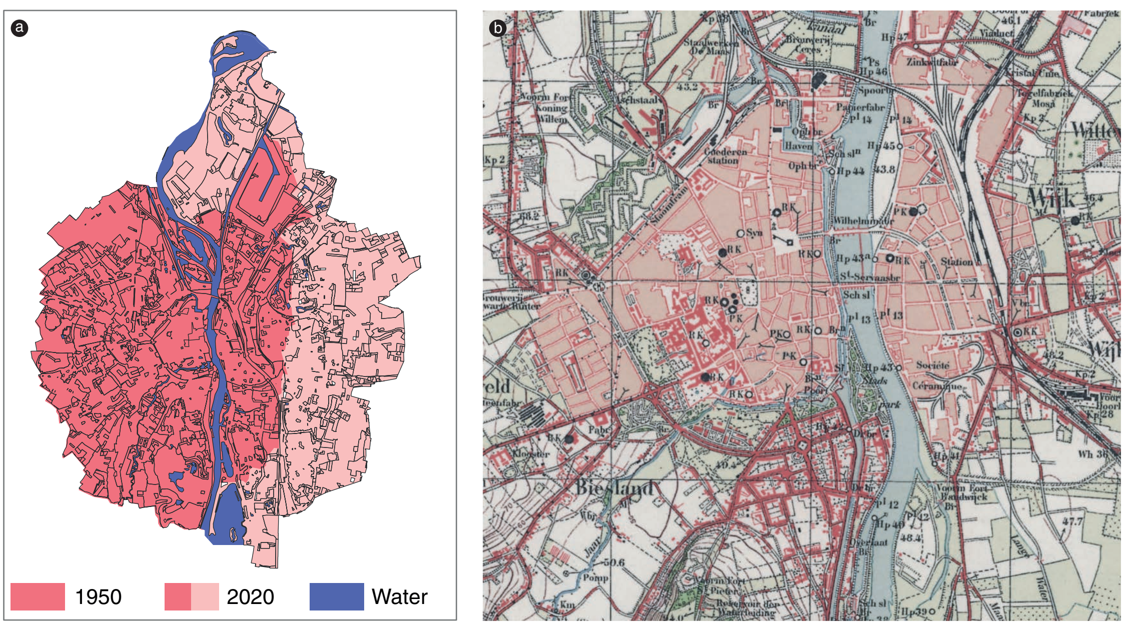


FIGURE 1.12 Some of the available data in the Maastricht case: (a) land use data derived from a topographic database – the 1950 and 2020 municipal boundaries are given for reference; (b) a detail from the 1950 topographic map to be digitized (Kadaster Geo Informatie)

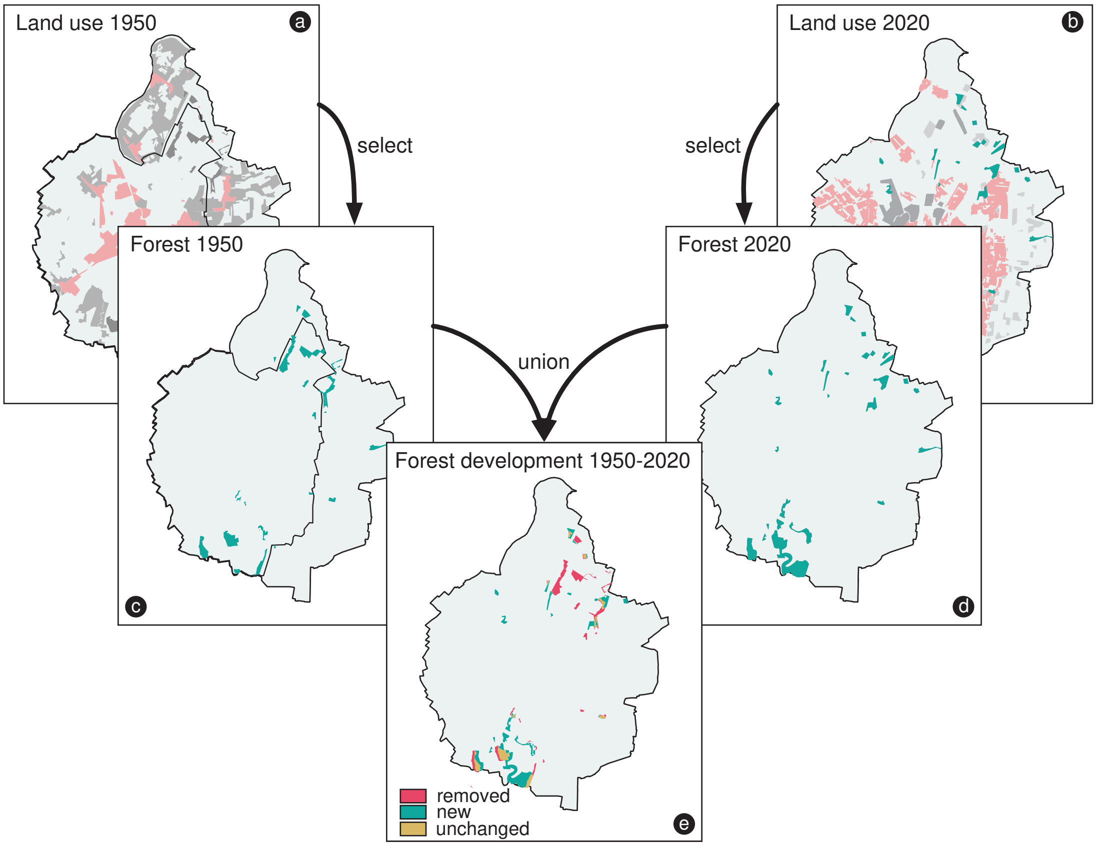


FIGURE 1.13 An overlay operation: (a) and (b) show the basic land use data from 1950 to 2020; (c) and (d) show the selected forest parcels from 1950 to 2020; (e) shows the overlay result: removed, unchanged and new forest parcels

expectations when there is some familiarity with the data. If the map revealed large new forest parcels in the city centre, it would be safe to conclude that something went wrong. If this were the case, one would have to correct certain steps in the analysis or refine analysis conditions.

Looking at the result in the map in Figure 1.13e, everything seems to be right. However, if one takes a closer look at the result, it can be seen that the result is not quite perfect. Figure 1.14a shows why. It reveals lots of small polygons indicating change. However, if the enlargement in the same figure, together with the two map details, is analysed (Figure 1.14a–c), it is obvious that in reality nothing has changed at all, although the GIS operation created 11 new polygons. These polygons are called 'sliver polygons'. A comparison of the basic statistics of the resultant data set with the original data set would have caused suspicion as well. The original data sets have 32 and 36 polygons, respectively, while the new data set has 100 mainly small polygons. The main reason for their occurrence is that the same feature in both the 1950 and 2020 data sets has a different geometry. The digitizing of the 1950 and 2020 maps was not performed by the same operator. Even if it had been the same operator, it would have been unlikely that the same points would have been selected during both digitizing sessions. For a problem like this, most GIS software packages offer simple solutions. One can delete all polygons with a size smaller than a certain threshold or calculate an average polygon boundary. Both approaches have disadvantages. But if the sliver polygons are not removed and the results are used in future geospatial analysis, the errors will propagate into the future results, and the database might grow unnecessarily.

During the collection of geospatial information, many types of error can be made: errors in measuring, classifying or categorizing data, localization errors, mistakes in data entry, etc. When these data are not directly incorporated into a GIS during the collection process, but are, for instance, mapped first because the new technology had not been applied yet, other kinds of error will emerge. Among these are generalization errors

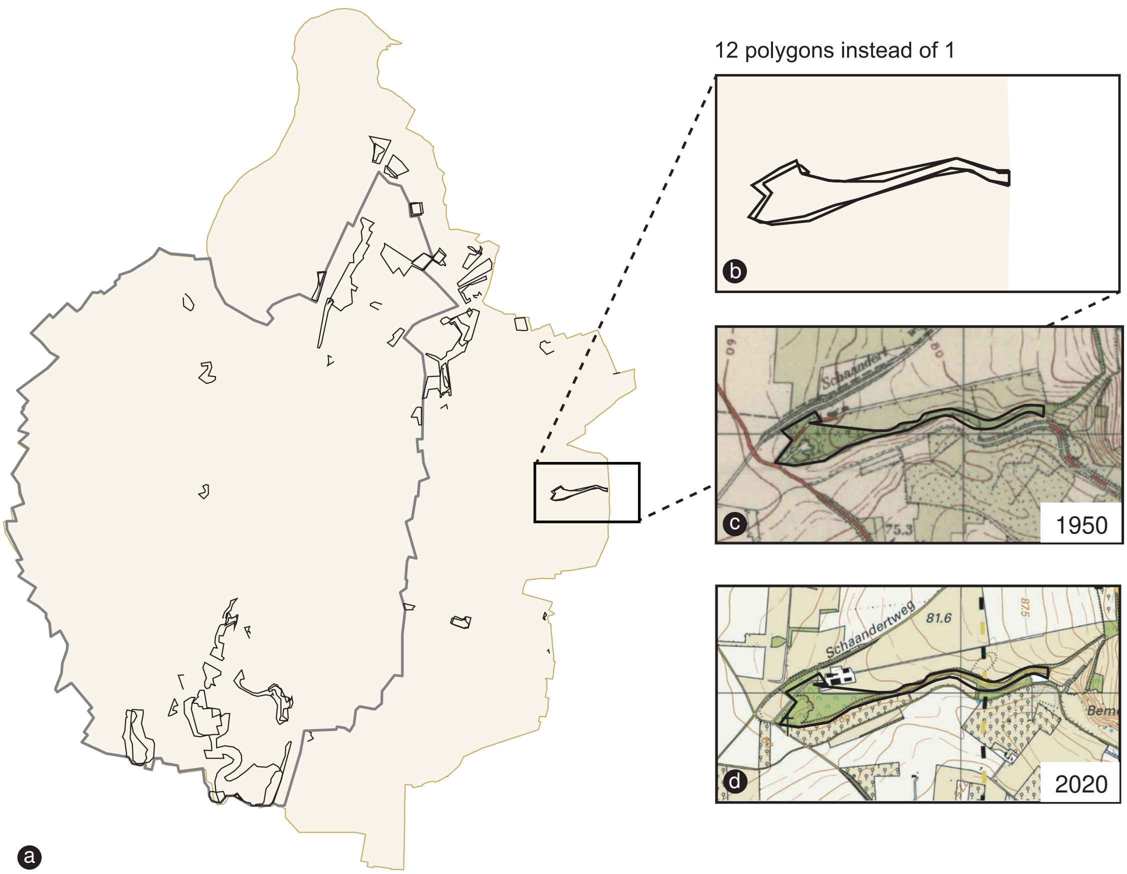


FIGURE 1.14 Overlay results and sliver polygons: (a) the forest parcels; (b) a detail: from 1 to 12 polygons; (c) and (d) the original 1950 parcels and the original 2020 parcels (Kadaster Geo Informatie)

or misrepresentations due to data amalgamation, reproduction errors and errors due to deformation of the printing paper. When these map data are subsequently digitized or scanned for input into the database, these errors are at least duplicated in it, but more probably the digitizing process itself will be another source of error as the Maastricht case demonstrates. At this moment, it is not quite clear, however, how these errors (for errors and their propagation, see Heuvelink, 2019) may affect the geospatial analysis results, i.e. whether they would lead to uncertainties in the results of analysis operations that would exceed some critical level. Not only are there errors in the input values, but also errors can be caused by analysis operations themselves and by the application-based models used. Examples are geospatial computational modelling techniques that forecast groundwater flow or polluted air diffusion and try to approximate reality but might in fact 'misrepresent' it. The combination of input error and these geospatial modelling techniques might lead to other error types. It is therefore very important to make sure that the data quality (i.e. suitability for specific applications – its fitness for use) is sufficient before basing decisions on maps that represent the results of geospatial operations executed on these data.

The results of a geospatial analysis operation are often presented in a report with maps, diagrams and tables to emphasize certain points or illustrate the conclusions. Most GIS packages do have a basic cartographic functionality to create the graphics. However, dedicated desktop packages have a more extended cartographic functionality and are better suited to producing the final maps. An example of these are the maps created by the municipality of Maastricht to illustrate its final report on the development of municipal forests. Some of these are shown in Figure 1.15. Included is a qualitative map that shows forest developments as well as a shaded topographic map with the forest development map draped over it to show relations with the terrain surface. Chapters 5 and 7 discuss the characteristics of the most existing map types.

It would be possible to execute the geospatial analysis discussed above without maps. For instance, one could ask for the area of forest stands in both years, compare them and answer the question without ever

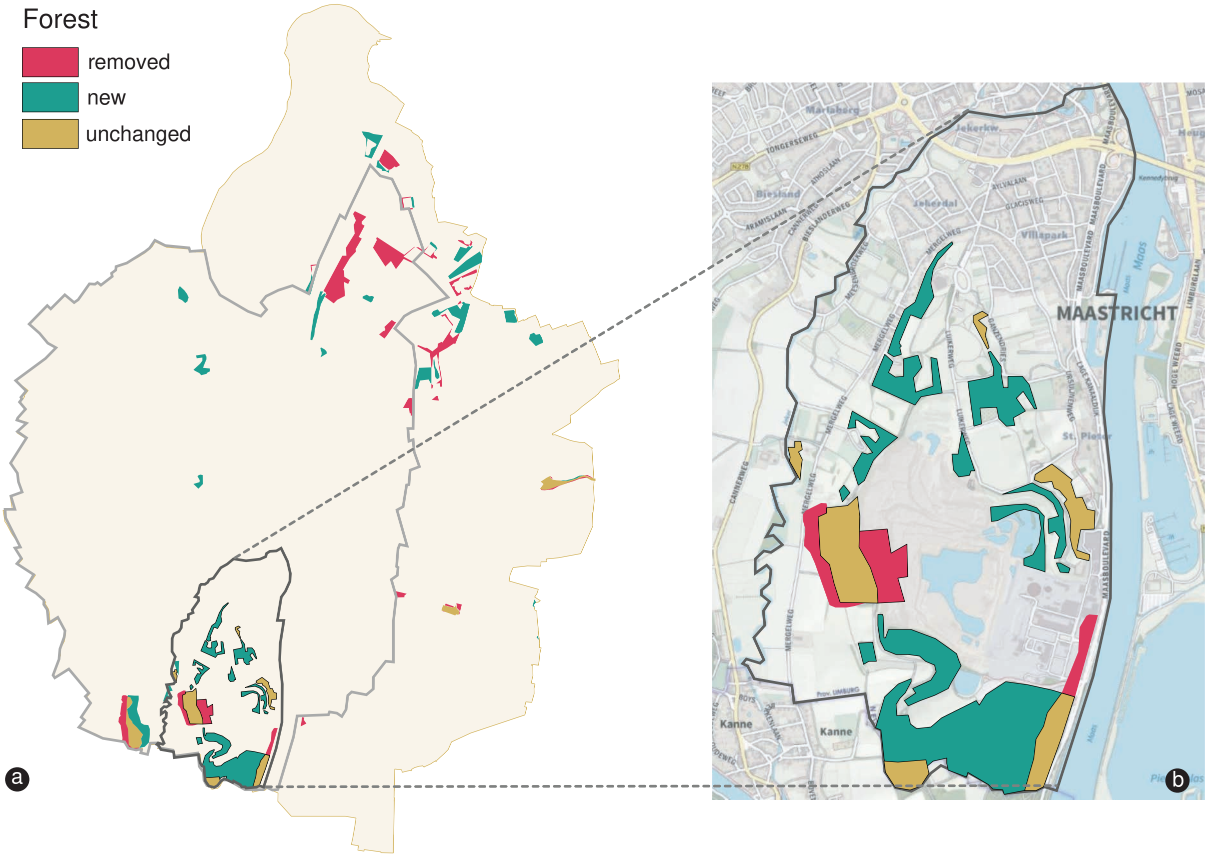


FIGURE 1.15 Presenting the results: (a) a qualitative display of forest developments in Maastricht; (b) relation between forest development and the terrain 1950–2020 (Source: J.W. van Aalst, www.opentopo.nl CC-BY)

visualizing them. But by doing so, one would deny oneself the opportunity of obtaining additional information from the process, as any geospatial trends or patterns in the answer may never come to light. The same size of forest stands in both years would not necessarily mean that they were at the same location. The information transfer without maps (e.g. through tables) would be more cumbersome as well. For instance, an increase is not necessarily due to the activities of the company active in the marl quarry, while the map would give a clear answer. It would be wise to apply the 'cartographic method', i.e. to visualize the geospatial relationships between the objects using abstraction techniques and the transformation based on the graphical grammar explained in Chapter 5; in other words, to map it. Observing the geospatial connections, relationships and patterns is only possible through the abstracting capacity of maps. As an example, another practical case study will be elaborated, showing the role that maps can play.

This case deals with the location of the TGV or high-speed train in the Netherlands 'Randstad' area when it was decided to extend the Paris–Brussels TGV link to Amsterdam. Consequently, as the existing rail links were already overburdened with traffic, and as extra foundations had to be constructed because of its high speed, a new route for the rail link had to be selected, through the 'Green Heart' of the Randstad area, i.e. the non-urbanized centre of the urban agglomeration in the western Netherlands. The proposed route should spare the environment as much as possible (nesting birds should be disturbed as little as possible, and there should be no polluting influence on groundwater or on vegetation). Moreover, no valuable geoscientific monuments should be affected.

An environmental information system (EIS) built for the Netherlands was therefore consulted. This EIS contains data on soils, groundwater, vegetation and fauna and even on rare geological outcrops (geoscientific monuments). These data had been collected for the EIS on a grid-cell basis. For each grid cell (1 × 1 km), dominant soil types (by putting the grid over a soil map), dominant vegetation types, the number of different vegetation types found per grid cell, the total number of vegetation type units, the types of wild animals that occurred, etc. were ascertained. Because of the fact that this information was stored in the EIS, the effect of the proposed routes could be easily estimated. In order to select the best route from a number of alternatives (Figure 1.16a), the susceptibility of the soils to water-table lowering (Figure 1.16b), the susceptibility of mammals to fragmentation of their habitat (Figure 1.16c) and the effects of disturbances and pollution on bird life (Figure 1.16d) for all the affected grid cells were determined. Subsequently, the computer was used to calculate how many of these grid cells would be affected, and to what degree, for every proposed route. In other words, one could use the computer to define the sum of the environmental values, which, because of the construction of the TGV along the various routes, would be affected or nullified. This created the opportunity to select the route that would create the least damage.

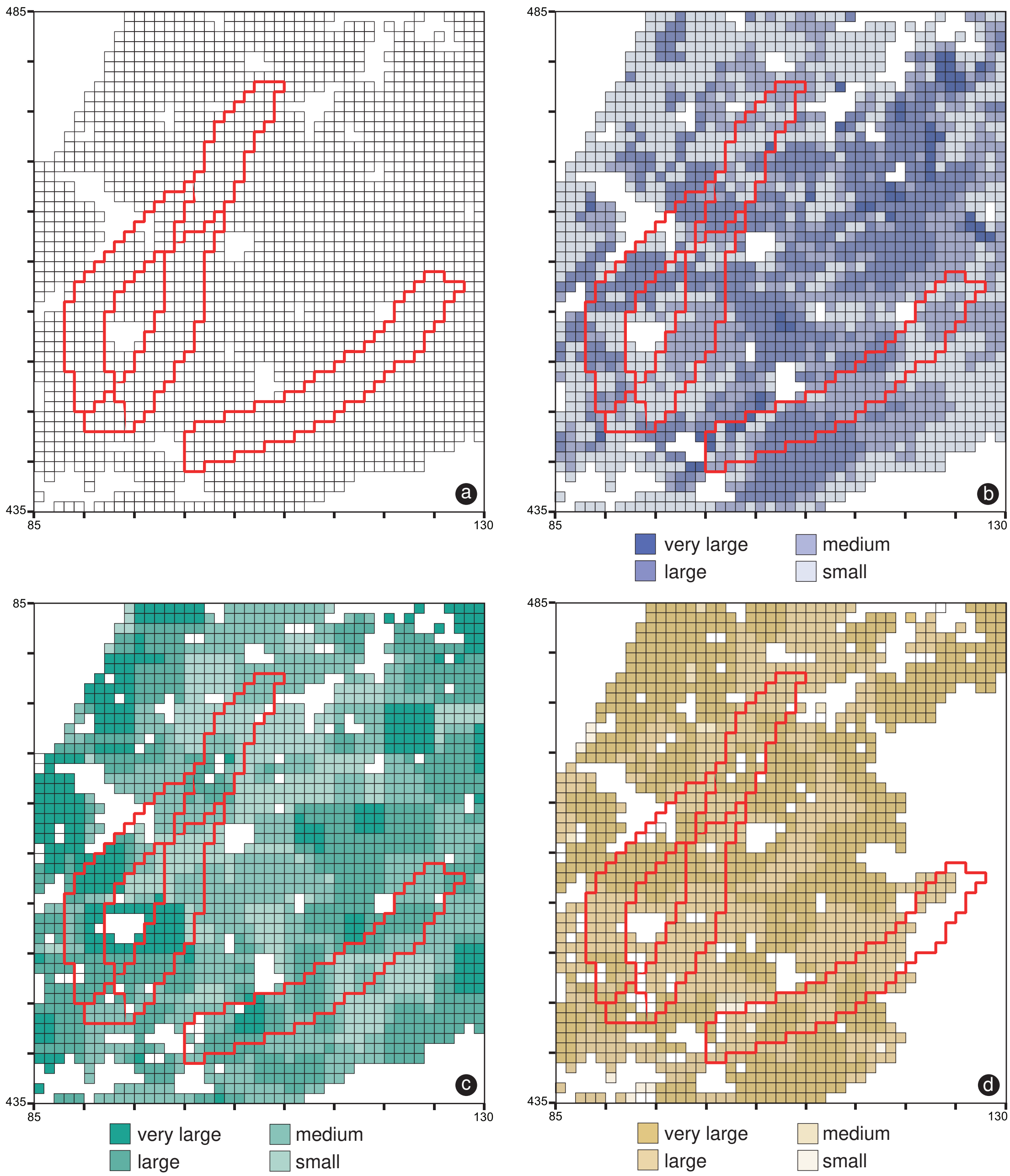


FIGURE 1.16 The location of the tracks of the TGV or high-speed train in the western part of the Netherlands: (a) alternative routes; (b) susceptibility to water-table lowering; (c) susceptibility of mammals to habitat fragmentation; (d) susceptibility of birds to traffic intensification

Digital maps are used outside of GIS as well, as will be discussed in Chapters 9 and 10. Electronic atlases are one example. Also called 'electronic atlas information systems', their function is less one of information processing than of answering specific questions, providing the support to integrate the answers in the mental map of the atlas user. This requires specific scenarios for a gradual immersion of the user into the new information environment. These atlas information systems can be extended to contain drawings, photographs, text and sound, and so become multimedia systems.

GISs are not yet well enough equipped to handle multitemporal information, and it is here that animated cartography comes in. Animation techniques are being developed that show the geospatial effects of developments at every stage. This presents extra potential for analysis and is one of the avenues for advanced data exploration that will, in future, also be available in a GIS. The potential for analysis is already greatly enhanced by the possibilities of applying GIS processing on the Web to all the data that have currently been made available there through the SDIs that have been constructed all over the world.

1.5 THE SPATIAL INFRASTRUCTURE AND MAPS

Companies and government departments at different levels (municipal, provincial/county, state or national) create and use information, for inventorying objects they administer or manage; for monitoring the state of the environment, or crime; for forecasting the weather, sales or schooling needs; or for reacting quickly to emergencies. In executing their day-to-day task, they offer part of this information to others and often need data from other organizations as well. Location is often the glue to link the different data sets.

Now this location component (street addresses, postcodes or ZIP codes, geographical coordinates or other geographical reference systems) allows one to combine the data from various information systems on the basis of common locations: it would be possible to link items from different files with the same postal address or coordinates. These possible linkages provide an enormous added value, but it would only be possible to realize this added value when a number of conditions have been met. Not only do the programs or packages used for the information systems have to be compatible, but also the structure of the files as well. When it comes to coordinate-based files, from which, for instance, maps of the phenomena could be created, joint ellipsoids and projection systems would be a necessity as well. Similar resolution is a prerequisite and preferably the data should be surveyed in a similar time frame as well. There seems to be little point in combining population data from the 1990 census with sales data from a 1999 sales drive.

The stimulus for spatial infrastructures originates from the above need to exchange data smoothly and is based on the motto 'collect once, use many times'. A GDI can be defined as (Groot and McLaughlin, 2000) 'A set of institutional, technical and economical arrangements, to enhance the availability (access and use) for correct, up-to-date, fit-for-purpose and integrated geoinformation, timely and at an affordable price, with the goals to support decision making processes related to countries' sustainable development'. In Europe, the GDI implementation is guided by the EU INSPIRE initiative, which, based on legislation, will implement the GDI concept throughout the European Union (http://inspire.jrc.ec.europa.eu/). In their guidelines, INSPIRE emphasizes several basic principles:

- Data stewardship and data security, meaning, e.g., that data should be collected once and maintained at the level where this can be done most effectively;
- Data accessibility and data interoperability must be possible to combine seamlessly spatial data from different sources and share it between many users and applications;
- Reusability and data synchronization are necessary: it must be possible for spatial data collected at one level to be shared between all different levels, e.g. detailed for those performing exhaustive investigations, but more general for strategic purposes;
- Data availability, e.g. spatial data needed for good governance at all levels, should be abundant and widely available under conditions that do not restrain its extensive use;
- Data discoverability, data validity and data rights, meaning that it must be easy to discover which spatial data is available that fits the needs for a particular use and under what conditions it can be acquired and used. By nature, an atlas has several relevant facilities for data discovery, such as an index for geographical names, a topical index and index maps;
- Data usability, meaning that spatial data must become easy to understand and interpret

because it can be visualized within the appropriate context and can be selected in a user-friendly way.

INSPIRE follows the standards established by the Open Geospatial Consortium (OGC – http://www.opengeospatial.org/). This is a non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial and location-based services and has companies, government agencies and universities as members. The standards and protocols describe interfaces and encodings that allow the so-called geoservices to operate. This means that data providers can offer their products in a standard way.

In order to enable data users to find out whether data sets from different sources can be combined at all, clearing houses are being set up. Such clearing houses provide metadata on data files. Metadata ('data on data') refers to information on the quality, time frame, accuracy, fitness for use, lineage and completeness of the data and on the way the data have been collected. The major function of metadata is that it allows us to check whether specific data are available, suitable and accessible. One of the possible functions of a clearing house is to check whether spatial data are not collected more than once, by different government institutions, for example. If data sets are found, the Open Geospatial standards guarantee the data can be used.

For many years, the Web (we will use this term instead of the 'World Wide Web' or 'Internet') has been the medium to acquire and disseminate geospatial data. Today's Web-induced revolution has even further increased the number of people involved in making maps. Should mapmaking via GIS still involve geo-professionals, the Web includes potentially everyone having access to this new medium to create maps. This has led to potential situations where the organizations offering maps via the Web will never know what their map products will look like, because the mapmakers on the other end have so many interactive options (to render reality through symbols that represent feature or phenomenon categories).

In the process of acquiring and disseminating geospatial data, maps are often used to represent the data. Maps can disclose or reveal geospatial patterns and help to increase insight into geospatial relationships. However, a specific design approach is required because of the nature of the Web (see Chapter 5), especially since the Web is highly interactive and users expect maps to be clickable. This interactivity and the possibility to use the Web to link distributed databases also make web maps good instruments to explore the different databases.

Maps can play different roles too (see Figure 1.17). They can function as an interface to other (geo)-information in cyberspace or can guide the surfer navigating the Web. In providing access by clicking objects, the map will bring the surfer to other web pages, which again can contain maps, photographs or text. In this respect, the map could play a prominent role in a country's SDI as being part of a search machine. In its role of guiding the surfer through parts of cyberspace, maps are used as a metaphor to keep track of paths through, for instance, a single website.

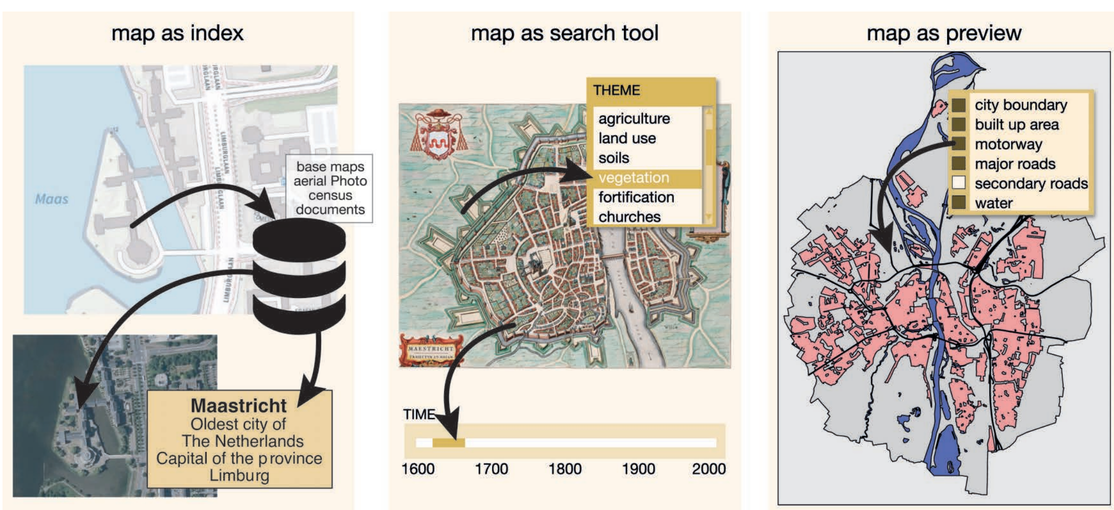


FIGURE 1.17 Maps used on the Web as index to other (geo)-information, as part of a search engine and preview of geospatial data offered

Why is the Web an interesting medium for maps? The answer is that information on the Web is virtually platform-independent. Also, many users can be reached at minimal costs and it is easy to update the maps frequently, although this last argument is only valid when the data provider is geared for it. Furthermore, the Web allows for a dynamic and interactive dissemination of geospatial data. This results in new mapping techniques as well as new possibilities for uses not seen before with traditional printed maps and most on-screen maps.

A true revolution was caused in 2005 by the introduction of the programs Google Earth and Google Maps. For everyone with Internet access, these programs were made available and could display satellite data and maps for free on a level of detail that was not heard of before. With an interface that is very intuitive in its operation, people could visit any location on Earth. Google Earth even allows three-dimensional flight through the landscape and cities. Not only could one visit places but it was also possible to add one's own data such as photographs, GPS (Global Positioning System) tracks and even maps to the Google environments to share these with others. This new but already very popular development was coined 'neo-geography' by Turner (2006). Figure 1.18 shows how Google Maps accommodates the sharing of collected photographs and other information. In Figure 1.19, it can be seen how users can drape their own maps on top of the Google Earth terrain and imagery data.

1.6 CARTOGRAPHIC EDUCATION

This book has been written to assist in cartographic education and intends, as a first objective, to provide an overview of the role that maps will play both today and in the near future in the world of geospatial data handling. To realize this, the nature of geospatial data is described as well as the characteristics of maps and the ways in which they can be put to use. Teaching map design is a second objective, which is shared out over topographic, statistical (or thematic) and temporal maps. Showing the manner in which maps function, either independently or combined in atlases, and can be analysed and interpreted is the third main objective.

The field of cartography is vast, and this book can only provide a selection of all subject matters relevant for the discipline. As such, it provides a subset of the Cartographic Body of Knowledge (BoK), i.e. all knowledge relevant for the art, science and technology of producing or using maps (a BoK presents the scope, boundaries and structure of a discipline, and it helps to establish expectations in professional practice and specialist skills (Fairbairn, 2017)). Currently, the International Cartographic Association (ICA) is engaged in the compilation of this Cartographic BoK. It has

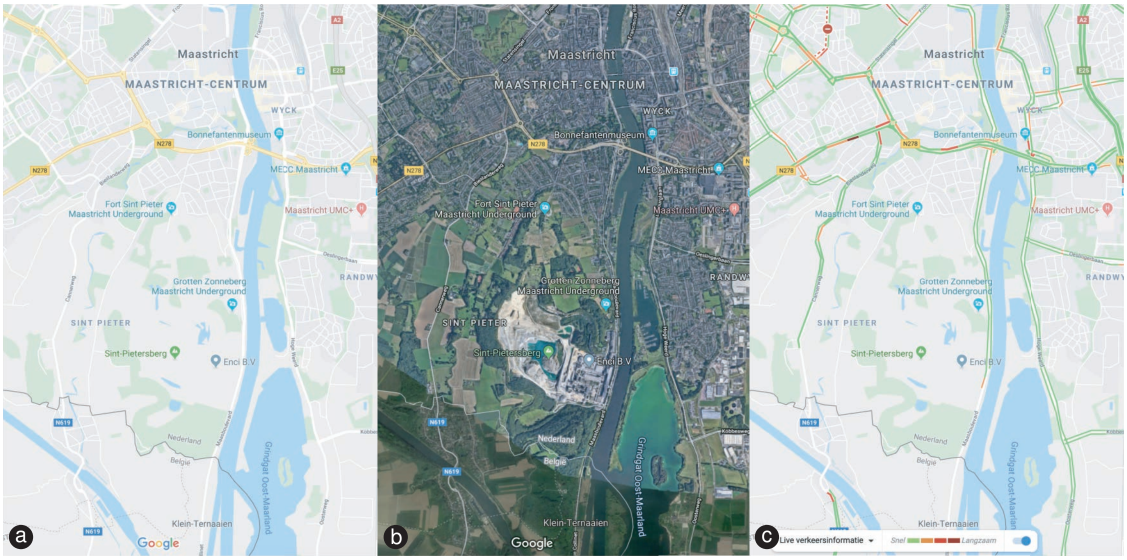


FIGURE 1.18 Google maps and several of its optional appearances: (a) topographic, (b) satellite image, (c) with traffic information (©2019 Google LLC, used with permission)

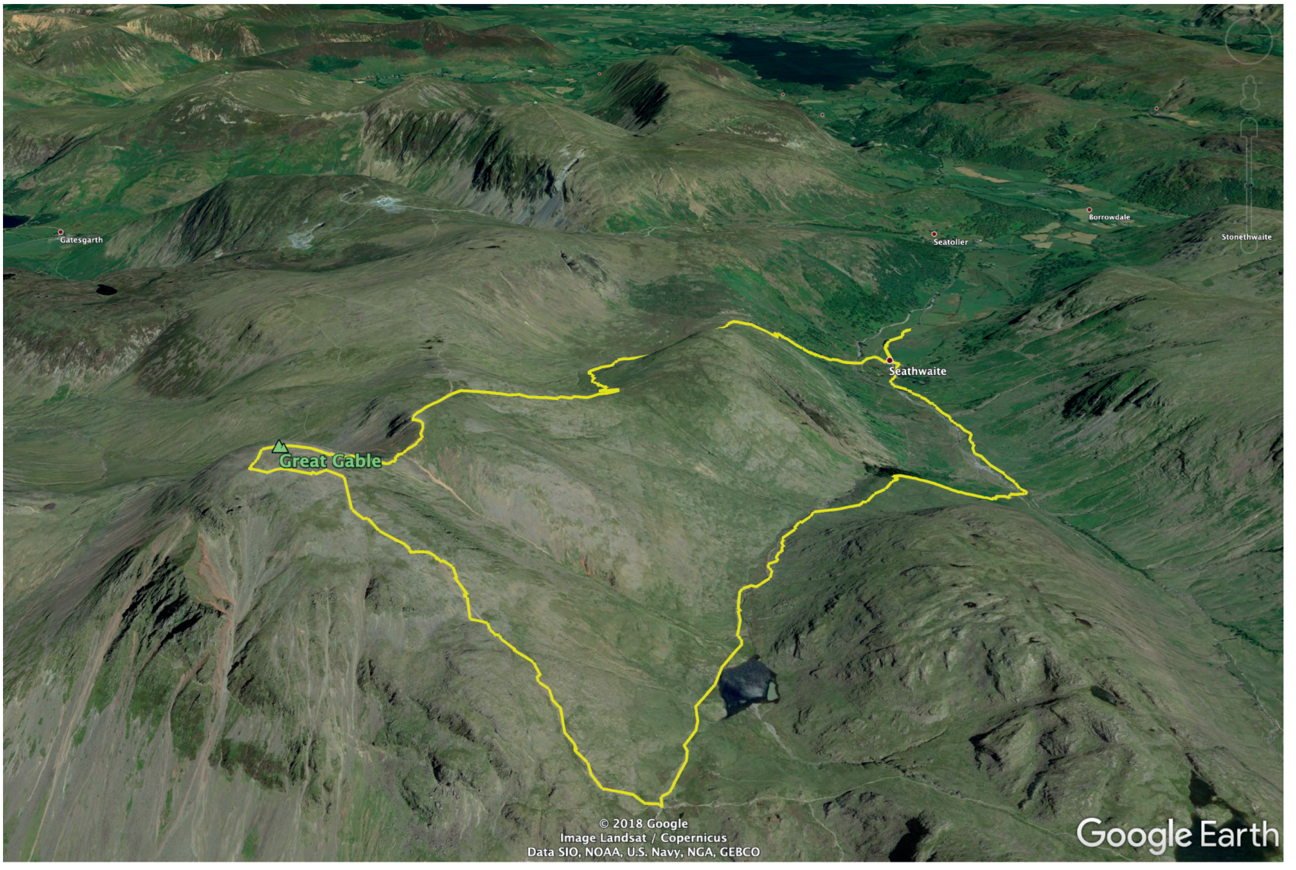


FIGURE 1.19 Google Earth: a three-dimensional view of part of the Lake District, south of Keswick and Derwent Water, with user-added data (©2019 Google LLC, used with permission)

instituted a special Working Group on Body of Knowledge on Cartography to realize this task. A closely related subject, Geographical Information Science and Technology, has already developed its own BoK, primarily directed to curriculum development (https://Gistbok.ucgis.org), and this BoK also contains a section on Cartography and Visualization, with learning objectives for some (2020) 32 subsections. For the time being, until ICA has completed its own BoK, this will serve as a well-informed substitute BoK for part of the field of cartography. The relevant ICA working group's endeavour is also based on the Strategic Plan of the International Cartographic Association 2019–2027, to be found https://icaci.org/files/documents/reference_docs/ICA_Strategic_Plan_2011-2019.pdf and on its research agenda for the coming years (https://icaci.org/files/documents/reference_docs/2009_ICA_ResearchAgenda.pdf). For those interested in pursuing a career in cartography, the items in this research agenda should be indicative for the direction we assume the field of cartography is moving in.

FURTHER READING

DiBiase, D. 2015. The nature of geographic information. An open geospatial textbook (E-education.psu.edu/natureofgeoinfo/) In State College: Department of Geography, Pennsylvania State University.

Groot, D., and J. McLaughlin, eds. 2000. *Geospatial data infrastructure – concepts, cases and good practice*. Oxford: Oxford University Press.

Heywood, I., S. Cornelius, and S. Carver. 1998. *An Introduction to Geographical Information Systems*. Harlow: Longman.

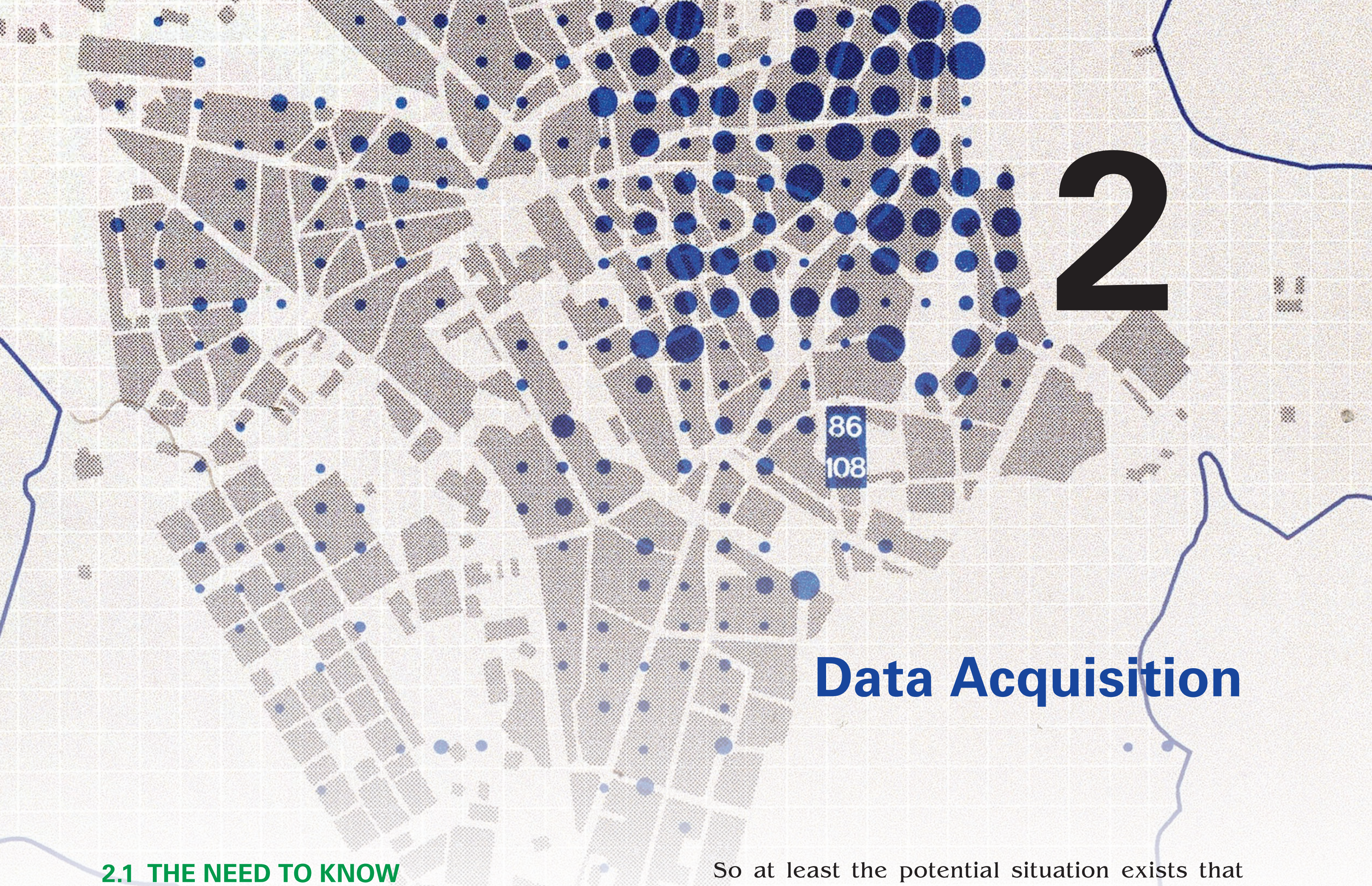
Kraak, M. J., and A. Brown, eds. 2000. *Web Cartography – Developments and Prospects*. London: Taylor & Francis.

Longley, P. A., M. F. Goodchild, D. J. Maguire, and D. W. Rhind. 2015. *Geographic Information Systems and Science*. 4th ed. Chichester: Wiley.

Rhind, D. 1998. *Framework for the World*. Cambridge: Geoinformation International.

Slocum, T. A., R. B. McMaster, F. C. Kessler, and H. H. Howard. 2008. *Thematic Cartography and Geovisualization*. 3th ed. Upper Saddle River, NJ: Pearson.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com



2

Data Acquisition

2.1 THE NEED TO KNOW ACQUISITION METHODS

In GISs, it is usual for many files to have been combined in order to boost the potential for analysis of the geospatial data. In an ideal situation, all the data combined have been collected, identified and measured on the same date, with the same geospatial resolution, according to identical procedures, and consecutively entered into the GIS using the same method. It is only then that the users would be sure of an adequate quality of the results of the analytical operations for which the files are being combined.

The reality is nowhere near this ideal: data are collected at different moments, are valid for different spans of time and have a different geospatial resolution; some might be collected in the field, while others were taken from existing maps that were generalized to an unknown degree. Some files might have been entered after they have been made compatible using some rubber-sheeting technique; others have been subjected to numerical transformations from other projections. Some might be based on random samples, while others on complete surveys. If it is numerical data collected at regular sample points, some might have been interpolated on the basis of linear, while others on the basis of geometric types of progression.

So at least the potential situation exists that the data, when compatible at face view, might not warrant the conclusions drawn from their analysis. It could be the case that the analyst or the GIS user in general should be warned about the results, for instance, that these should be interpreted with care. Traditional topographic maps used documentation helps, like reliability diagrams to show that navigation in certain areas on the basis of the map would be hazardous, as the producer could guarantee neither the accuracy nor the completeness of the data. In the more complex world of GIS, one needs numerical aids to be informed about the quality of the data files, in order to be able to decide on the validity of analysis results.

It should be kept in mind that in all these cases, one is collecting coordinates with which to describe locations of objects, with attributes the nature of which is determined either at the same time (e.g. during terrestrial topographic surveys) or later, in a laboratory (as is the case, e.g., for soil surveys), or through field checking (as is the case for remotely sensed data).

The various geospatial data acquisition methods used in GIScience (geographical information science) can be divided into the following types (see also Figure 2.1).

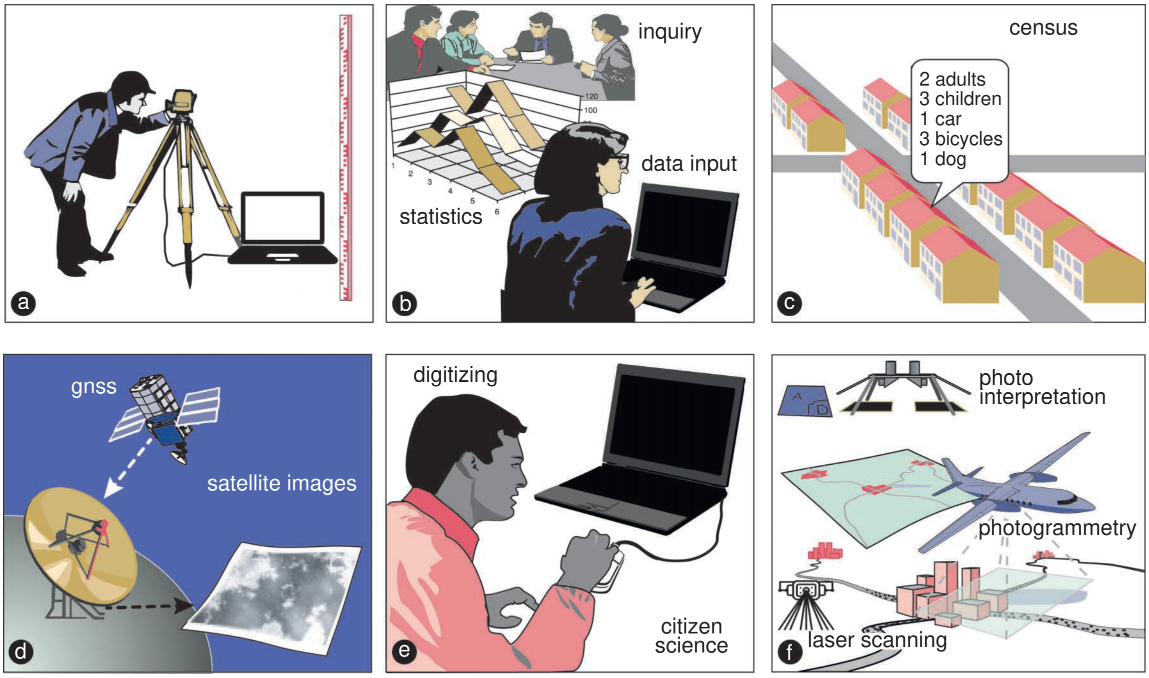


FIGURE 2.1 Various types of data acquisition methods: (a) surveying; (b) enquiries and statistics; (c) collecting census data; (d) remote sensing – GNSS (Global Navigation Satellite Systems); (e) digitizing maps – citizen science; (f) photogrammetry and laser scanning

2.1.1 Terrestrial Surveys

Large-scale topographic data can be acquired through terrestrial surveys. Such surveys immediately lead to digital files that can be imported into a GIS. When surveying new extensions for telephone companies' or cable companies' networks that dispose over digital files of their networks, surveyors would use 'total' stations with which the survey data are immediately edited and transformed into files that are extensions of existing files, and can be added to these main files at will.

2.1.2 Photogrammetrical Surveys

From aerial photographs produced from manned aircrafts or close-range cameras (drones), object coordinates can be determined in digital stereoplotters and imported directly into information systems. The attribute information required could be determined either through interpretation or through field checking.

2.1.3 Lidar or Laser Altimetry

Lidar or laser scanning equipment sends out a beam of light which is returned by the surface it touches upon. The result is a large point cloud of data, which after corrections can be used to reconstruct the objects or the terrain. The scanners can be mounted on a tripod, a backpack or car for terrestrial surveys, or on plane, helicopter or drone for airborne surveys. GPS (Global Positioning System) equipment at all times keeps track of the location of the scanner to be able to determine the accuracy of the measurements that are within centimetre range.

2.1.4 Satellite Data

Satellites have been built that contain scanners with sensors susceptible to radiation emitted or reflected by the Earth's surface. These scanners operate in such a way that the sensors measure radiation sequentially from patches or grid cells along paths perpendicular to the line of flight. These radiation data are later put in their proper geospatial relationship, and by doing so, a map is simulated. Here, the data accuracy also depends on a number of correction techniques, both for the radiation values and for the geometric accuracy. After these corrections, the data frequently have to be resampled in order to fit specific grids, or to be comparable with other data sets. By being collected for grid cells (with, for instance, 5×5 or 20×20 m resolution), the data are generalized from the start.

2.1.5 GPS Data

A special set of satellite data is provided by the GPS, which, on the basis of 24 satellites, is able to pinpoint one's position three dimensionally with an accuracy of some centimetres. GPS is the name of the positioning system developed by the United States. In 2020, GPSs developed by the European Union (Galileo), Russia (GLONASS) and China (BeiDou) are also operational. GPS will become a generic term for positioning systems. These GPS recordings are used to increase the accuracy of existing georeferencing methods, or can be used directly in data surveys; they can be used for both point surveying and linear surveying. The data are recorded on the basis of a global reference system, which can be transformed to a local reference system.

2.1.6 Digitizing or Scanning Analogue Maps

Digitizing refers to the conversion of analogue images into digital representation. Initially, this was done manually by registering with cursor sequences of characteristic points belonging to lines on a map. Through this action, the coordinates of the positions touched were recorded digitally. Nowadays this conversion is mostly effectuated through scanning (see Section 2.4.3): optical records of the existence of specific colours at specific positions on a sheet are transformed into files with information on positions with attributes (hue or colour value). If digitizing or scanning could be effectuated with 100% correctness, the results would still depend on the accuracy of the original maps.

2.1.7 Using Existing Boundary Files

From open data domains, governmental data portals and commercial organizations boundary files (digital geometric descriptions of administrative units) can be acquired for applications that occur more frequently or for a larger number of uses: topographical files, files for car navigation, boundary files to be used in conjunction with statistical data (like for marketing applications). For these existing files, it is essential that they are compatible with one's software and therefore are based on the same standards as used by the buyer.

2.1.8 Socio-economic Statistical Files

National statistical services are increasingly publishing their (census) data online, with the appropriate software to query the files or even to visualize them in map form. These files are mostly presented in standard formats that make them compatible with most current mapping packages. One of the items contained in these socio-economic data files is the standard area code that relates the data to the appropriate areas for which they have been collected. This link between data and area should also be preserved when entering the data into other packages. It is through these code links that these packages will know what data to link to specific areas (see also Figure 6.48). On a global level, the United Nations Sustainable Development Goals indicator database provides current statistics (https://unstats.un.org/sdgs/indicators/database/).

2.1.9 (Geo)physical Data Files

Earth scientists have been producing data files since the late 1980s and were among the first to make them available to the public at large. It was to provide base maps for these global physical data files that the Digital Chart of the World (DCW) (see below) was first conceived (a base map is either a map showing topography that enables one to understand the location or distribution of a phenomenon (such as in Figure 3.7) or the (relatively) large-scale topographic map of an area from which all other, smaller-scale maps of that area are derived (such as in Figure 3.3)).

2.1.10 Environmental Data Files

Some examples are the Data Distribution Centre (DCC http://www.ipcc-data.org) of the Intergovernmental Panel on Climate Change (IPCC), which provides climate, socio-economic and environmental data, and the Corine Land Cover data from the European Union (https://land.copernicus.eu/pan-european/corine-land-cover).

2.1.11 Volunteered Geographical Information; Citizen Science

As fieldwork becomes increasingly expensive, mapping or nature conservation agencies become more and more dependent upon the information received from citizens that report on, e.g., changes in the topography or toponymy, and on the number of bird or vegetation species counted or found. For areas without detailed topographic maps, in emergency situations, volunteers might develop plans and maps during mapathons, sessions organized under the Missing Maps formula, to provide humanitarian agencies with the maps they need, on the basis of current satellite data and the information from local informants.

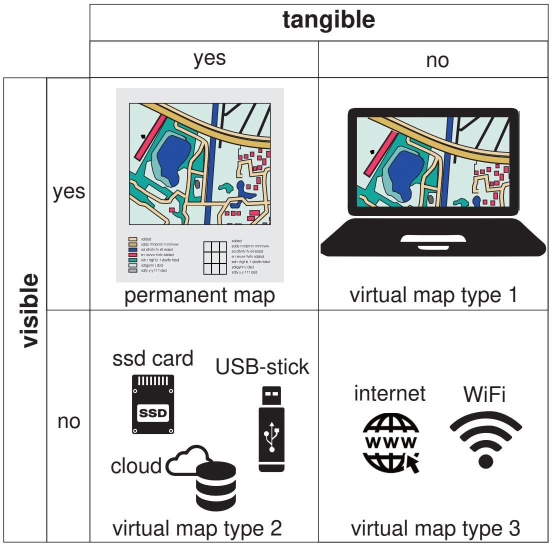


FIGURE 2.2 Analogue and virtual maps (after Moellering, 1983)

Before they are entered into the GIS, data from these sources are stored in different ways: some in the form of paper maps and some in the form of files (e.g. commercially produced files well protected by copyright (like boundary files) and published on online; remote sensing files; socio-economic, environmental or (geo)physical files; files of oceanographic surveys; photogrammetrically plotted data; or GPS data that are distributed mainly through the geodata infrastructure portals).

It is in relation to the various types of storage or procuring of geospatial data that one discerns between analogue maps and various types of virtual maps (after Moellering, 1983). Virtual maps of the first type can be seen, not touched: these are the maps made visible on monitor screens. Virtual maps that are not visible but tangible are the ones that can be downloaded from the Internet. These are dubbed virtual maps type 2. The third type of virtual maps is neither visible nor tangible and can be procured through Internet. As soon as one has queried them from the Internet, they can be viewed on a monitor screen (type 1) or printed as analogue maps (Figure 2.2). The virtual maps type 3 can be equated to the digital cartographic model (see Sections 1.2 and 2.5.1).

2.2 VECTOR FILE CHARACTERISTICS

Though actual file structures will be dealt with in Section 2.5, Sections 2.2 and 2.3 will focus on other aspects of files used as base maps (boundary files) for GIS maps that are available via portals. Before doing so, vector files and raster files have to be differentiated. In vector files, lines or boundaries between areas are defined by a series of point locations and their connecting links; in raster files, boundaries, or any other relevant background information, are defined as strings of picture elements (pixels) in regular grids that have been activated with specific values (see Figure 2.3).

To use these vector files as geographical reference frames for the geospatial information one wants to visualize on a monitor screen, one has to take account of the following aspects: resolution (the relation between the area as represented by a pixel and the same area in reality; see Section 2.3), digital scale and the possibility to separate the files into different categories or layers.

Vector files, to be used as boundaries, can be produced by scanning maps (see Section 2.4). This procedure concerns maps of a specific scale – of course, once the information is digitally available (through digitizing), scale seems to become less important issue, as it will be possible to zoom in or out and thus change the scale at will. But whenever the scale is increased beyond the original scale, the ensuing image will look poor and coarse, and cannot be used further for precise referencing. This is because the original map that was digitized will have been subjected to generalization (see Section 6.4). Apart from decreasing the scale, the projection of the map can be changed easily; as soon as the proper transformation formula from one projection to another has been determined, the digitized data can be displayed in any other projection system.

An early global product was the DCW (also called 'VMap0') (a data file containing all the linear elements (coastlines, rivers, contour lines, state boundaries, major roads and railways and cities) was produced mainly by the US military mapping agency National Geospatial-Intelligence Agency (NGA) in 1990 from Operational Navigation Charts (ONCs) at the scale 1:1 million). These ONCs are ultimately based in most cases on topographic maps 1:50 000 or 1:250 000, produced from aerial photographs. Both in producing the 1:50 000 maps from the photographs and in producing the 1:1 million charts from the 1:50 000 maps, generalization has been applied with its ensuing simplification, exaggeration, displacement and selection (see Chapter 6 for generalization). Objects that would still be retained at a scale 1:100 000 or 1:250 000 are omitted on the scale 1:1 million. So when the DCW is zoomed in on, and enlarged to a scale 1:100 000, it would not show all objects one would expect on this scale.

It would be otherwise if large objects and minor objects were stored in different layers or files and

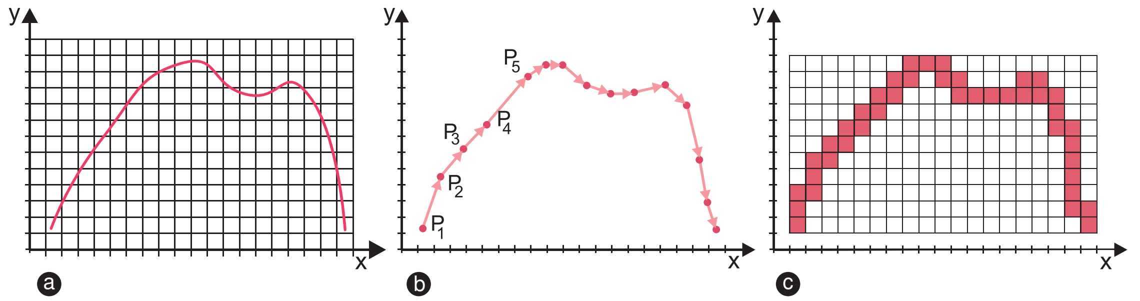


FIGURE 2.3 Representation of a line (a) in vector format (b) or raster format (c)

could be activated whenever a specific threshold scale value was passed: by zooming in beyond the scale 1:250 000, minor rivers could be activated and be made visible, etc. So it is important that (boundary) vector files have their objects stored in at least as many layers from which a selection can be made for display. An example of such a subdivision of objects into categories would be:

Hydrography: (a) Major rivers, lakes; (b) minor rivers, lakes.
Territorial boundaries: (a) National boundaries; (b) state/provincial boundaries; (c) county boundaries.
Coastlines: (a) Coastlines and major islands; (b) minor islands.
Administrative names: (a) Names of countries; (b) names of states/provinces; (c) names of counties/departments.

The various administrative areas should be supplied with the codes assigned to them, for the purpose of being able to match them to the statistical files (see Section 2.1).

This approach is currently followed by global map products such as OpenStreetMap, Google Maps and Bing Maps, which have organized their maps in 20 tiled layers, each with content appropriate for that particular scale level. Tiles are pre-rendered maps which speed up the map rendering process of user's area of interest. However, most of these are raster tiles and only useful as base map and not as layer to be used in a spatial analysis.

The DCW project has been replaced by Global Map, an initiative led since 1996 by Japan, to make available topographic files on a country basis with a similar resolution as DCW but more recent data and a more extensive thematic coverage. These geographical data sets are composed of the following thematic layers: elevation, vegetation, land cover, land use, transportation, drainage systems, boundaries and population centres. The major reason to develop this global data set was to enable the monitoring of changes in the global environment. After being transferred to the United Nations, the project ended in 2018. The original DCW files can still be downloaded from the Pennsylvania State University library site (https://psu.box.com/v/dcw) or be obtained from ESRI, be it in an updated form.

Other examples of vector files are EuroGeographics' EuroBoundaryMap files (see Section 6.6.2) and the Electronic Chart Display and Information System (ECDIS). Member states of the International Hydrographic Organization (IHO) developed regional databases for ECDIS (using a Digital Data Transfer Standard (S-57), with a common object code and exchange format (DX-90)). Though preference is given to the production of regional databases covering all the routes used by international shipping, this will take time. Such data has not been generally available worldwide before 2002. In the meantime, hydrographic chart raster files, like the one developed by the United Kingdom, will be used.

2.3 RASTER FILE CHARACTERISTICS

In a scanner, optical records of the existence of specific dots, dashes or patches in black and white or colour at specific positions are transformed into files with information on positions with attributes. As the registration of these dots, dashes and patches takes place along regular parallel scanning paths, in incremental temporal steps, the output of these files consists of regular grids, built up from picture elements (pixels), each representing specific attributes (see Figure 2.3c). The higher the resolution of the scanning device (and its price), the smaller the pixels will be, and the more detail of the original image will be rendered.

To revert to the case at the end of Section 2.2, existing hydrographic charts are now being scanned and stored digitally, to provide charts in a form that can already be used in ECDIS, in order to bridge the gap before they will be made available in vector format. The first to do so was the British Hydrographic Office, with its ARCS (Admiralty Raster Chart Service) package. ARCS consists of DVD files produced by scanning existing charts, with an update option that allows one to combine data from weekly DVD notices to mariners with the original raster chart to produce new charts on screen that are fully corrected. Figure 2.4. shows a detail of the OpenNauticalChart (http://opennauticalchart.org/).

The resulting charts when displayed on the monitor screen can be overlaid with radar images and also visualize the ship's position as well as data on its course, speed and planned track. The electronic navigation systems using ARCS DVDs will allow for the selection of the area to be viewed, and also allow for zooming in or out and for adding or omitting a number of additional data layers. When the information contents of a paper map are available on different films that can be scanned separately, the raster file can consist of different layers – in this case with contour lines, names, buoys, bathymetric colours, etc. that can be activated either separately or together.

Apart from updating, the raster image itself cannot be queried or otherwise electronically analysed – it forms a backdrop picture against which geospatial processes can be visualized. Important aspects of these backdrop raster files are the same as for the vector files: the scales at which they have been scanned and the existence of information in different layers.

Another example of such a raster file are the raster versions of the topographic map series of the many countries. These files, available online, contain all the information of the paper topographic maps. These maps can be used as georeferenced background in GIS or in apps on mobile devices for leisure purposes. It is for products like

FIGURE 2.4 OpenNauticalChart (http://OpenNauticalChart.org/ (CC-SA))

this one that the resolution is important: scanners (see Section 2.4) register with a specific number of dots per inch, indicating the smallest areal units for which information is being determined. A resolution of 250 dpi (dots per inch), e.g., refers to the fact that areas with a size of 0.01 mm^2 are represented independently (see Figure 2.5). Equally important as the geometrical resolution is the radiometric one. It refers to the number of colours that can be differentiated between by the scanner and to its display capabilities as well.

In contrast to vector files, transforming raster files to other projections is very difficult. Coordinate systems with which the images are overlaid cannot easily be changed for other systems. Additionally, the pixel is the basic unit of the image structure, and it might not be relevant as a reference unit for the theme mapped. A pixel is a representation of the smallest area where electromagnetic radiation is collected from individually. The larger these areas, the less information there will be to store and the easier it will be to handle the resultant files. On the other hand, the smaller these areas on the Earth's surface, the more accurately the resulting raster image will model the original data.

Satellite imagery is made available as raster files. Resolution here is not a function of the scanner measurement device's size but of the sensor's field of view: by collecting the radiation within a specific angle and from a specific distance from the Earth, the sensor registers and measures radiation from a specific patch of the Earth's surface. The size of this patch, the dominant radiation of which is being registered, will determine the usefulness of the imagery for constituting a model of the objects whose radiation values are being registered. These element areas on Earth, represented as pixels on the satellite imagery, have decreased in size since the first civil satellites began sending their information to the Earth in the 1970s: Landsat registered 80 × 80 m^2 first and later (Landsat Thematic Mapper) 30×30 m^2; SPOT in 1986 20×20 m and even 10×10 m for its 'panchromatic' applications. Today, resolution goes beyond the one-metre resolution. Many countries as well as commercial organizations have satellite programmes. Some of these offer the imagery for free like the Copernicus programme of the European Union with their Sentinel satellite family. Up-to-date information on satellites, and their sensors and resolution, orbit height and repeat cycles can be found at ITC's satellite sensor database: https://www.itc.nl/research/research-facilities/labs-resources/satellite-sensor-database/.

When one compares raster files with vector files, the latter have a greater overall resolution. This is

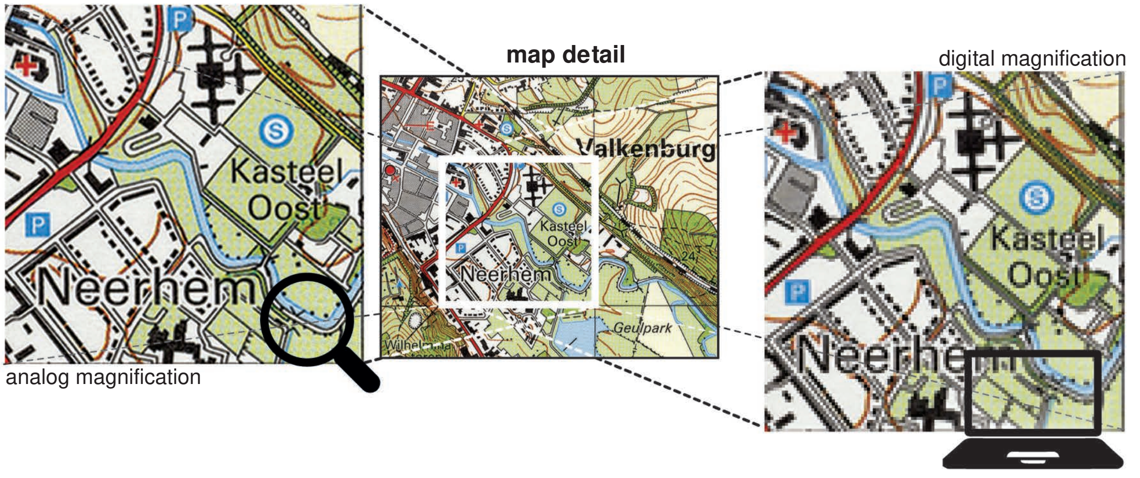


FIGURE 2.5 Tenfold enlargement of photographs taken from Top25 Raster (the raster-based screen representation of the topographic map of the Netherlands at scale 1:25 000 (right)) next to a tenfold enlargement of the analogue version of this map (left) (Kadaster Geo Informatie)

because in the raster technology, the input device provides the information divided into discrete pieces ('pixels') with a finite size. When the pixels are large, much information is lost. On the other hand, this grid structure of satellite and scanned images allows for analytical operations that are much easier and much less time-consuming than is valid for vector images. It also provides for relatively easy combination with other files, as well as for a multitude of image processing possibilities. As raster technology is also behind the monitor screens, vector images are simulated in reality on these screens, being built up from activated raster cells.

2.4 DERIVING DATA FROM EXISTING MAPS

The technical procedures for deriving data from existing maps will be described in Chapter 5. In this section, the requirements that existing maps have to answer to, in order to be suitable sources for digital files, will be indicated. In particular, it is documentary aspects that are important here. Such a plethora of paper maps has been produced that finding the proper one can sometimes be extremely difficult. Furthermore, the organization and aims will be covered, while the principles of the hardware used will be described. The data to be derived from existing maps are complex in the sense that they have both locations and (other) attributes. So as a first requirement, these locations and attributes should be unambiguous on the source documents: the definitions used or implied in the map legend should be clear and consistent, the period of time during which the data were gathered should be mentioned (in order to be copied and stored somewhere in the digital files as well). The data quality aspect is important and will be covered in the next section.

2.4.1 Finding the Proper Map: Documentation

Finding the proper map will depend on the selection of a map of the proper area, the proper theme and the proper time period in which the data have been collected. Most map libraries will contain map descriptions in their catalogues that refer to the area and theme of a map, and also have data on the map's date of publication. They will rarely have information in the catalogue files on the time period, in which the data for the maps have been collected, i.e. of the period for which the map is valid. For topographical maps, this is used to be about 2 years before the map has been published; for thematic maps, this might take longer.

2.4.2 Preparation

Finding out about the way in which locations have been indicated on the paper map is essential, as this will be a guide to the manner in which the information to be digitized or scanned will have to be transformed in order to fit in one's GIS. When preparing an existing map for scanning, the method according to which the data are geospatially referenced should be taken account of. This georeferencing (providing a geospatial address) can have been done by using either geographical coordinates, grid coordinates or no coordinates at all. In the latter case, the map cannot be further fitted into or integrated with existing files, unless by rubber sheeting (see Section 6.3). Using geographical coordinates, the Earth is regarded as a sphere; however, using grid coordinates, the area represented is regarded as a flat plane (see Section 6.1). The UTM (Universal Transverse Mercator) grid coordinate system combines the latter two in a sense, as it consists of a number of planes that together cover the Earth (see Section 6.2).

If, e.g., a file digitized at scale 1:250 000 of North-west Europe has to be extended or updated, topographic maps 1:50 000 of the Netherlands have to be digitized. This means that all coordinates implied in the file produced when digitizing the map will have to be transformed from the stereographic azimuthal projection in which the maps of the Netherlands are rendered to the UTM projection of the 1:250 000 file. Digital files are scaleless in principle, but their resolution is still governed by the scale at which the data have been digitized. Now, in order not to have discrepancies in the resolution, for the data from these maps of the Netherlands to fit in the new 1:250 000 North-west Europe file, they have to be generalized as well (see Figure 2.6).

2.4.3 Digitizing

As indicated in Section 2.1, digitizing refers to the conversion of analogue images into digital representation. The main means of converting analogue data into digital data are scanners. Before scanners were used, manual digitizers were employed, consisting of a tablet in which wires were embedded, located along Cartesian axes. A cursor was linked to the tablet and, when a cursor button was pushed somewhere on the tablet, the electrical charge generated was picked up by those wires directly underneath in the tablet, and the wires activated would then provide their specific codes as *x*- and *y*-coordinates. Additional buttons on the cursor would allow one to join attribute information to the locational information. For entering topological information (information about logical relationships between geospatial objects, such as about relative positions, adjacency and connectivity) such as stating whether a node belongs to a specific line, or is the beginning or end point of a line or is located somewhere in between, specific areas (called a 'menu') on the digitizing tablet could be activated and used as legend boxes: by pressing these areas by hand or button, the link between the digitized location and specific attribute information was forged and entered into the computer memory.

This hardware is hardly used any more, but it still serves well in order to discuss what the main characteristics are one should want to hold on to during this process of converting geospatial information from analogue into computer-readable form. It is to preserve the relationships that were visualized on the map. This means, e.g., (as shown in Figure 2.7) that

- Existing links between points on the map should be retained in the digital file;
- Parallel lines should remain parallel;
- Relative locations should be preserved;
- Absolute locations (as expressed in coordinates) should be preserved;
- Adjacency should be preserved;
- Lines that merely touch each other should not intersect each other.

Many digital files in current use have still been based on manual digitizing, and this might have resulted in the kind of mistakes as described above. These are avoided by using scanners (see Figure 2.8a). The operating principle of high-resolution scanners can be that of a rotating drum to which a paper

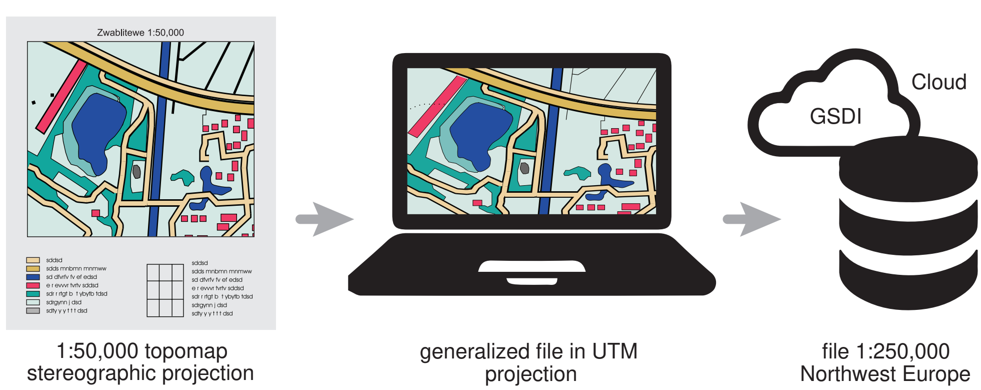


FIGURE 2.6 Conversion of an analogue map into a digital file

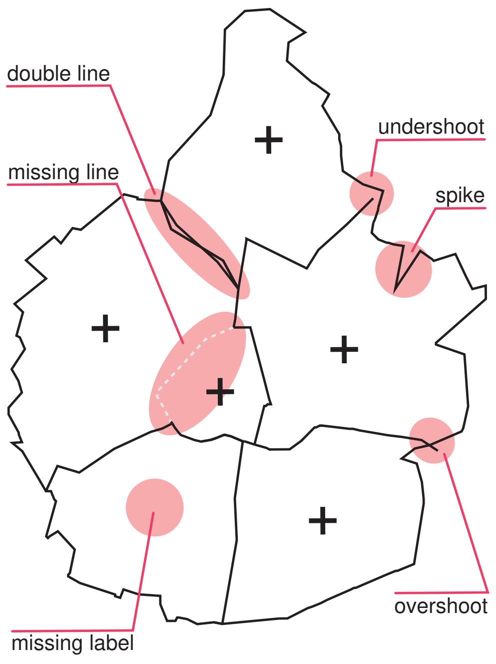


FIGURE 2.7 Potential errors as a result of manual digitizing

map has been attached. While rotating, a sensor will scan the map in narrow (e.g. 0.1 mm wide) contiguous bands and will register the light intensity – or colour value – of small squares (0.01 mm²). These light measurements are then transformed to digital values, and can be digitally represented on screen, thus reconstituting the scanned original. By splicing the signal picked up by the sensor over measurement devices susceptible to red, green or blue light, the original colours can be reconstituted.

The working principle of low -to medium-resolution table scanners used in desktop environments is that of registering the characteristics of an image on a page, put facedown on an A4 or A3 tablet, line by line (Figure 2.8b). All data points or pixels on a line are registered sequentially before moving to the next line. It subdivides images into discrete data points in which the sensor (ccd = charge-coupled device = electronic light sensor) measures their light value or colour value. In this way, an analogue map is turned into a description consisting of geospatial addresses (grid cells) and their characteristics (light values).

This is the easy part of scanning. After scanning, there is still the operation to make sense of the scanned data, and to change it back into a vector file, if necessary. If possible, it is not the printed maps themselves that are scanned, but the original colour separates prepared for the map's reproduction. On these, there is already a separation of functions that should then be further edited. On scans from the separates to be printed in blue, one should indicate whether lines refer to rivers, coastlines or lake shores; codes can be added to the rivers that should indicate their importance; this is relevant should the scanned image subsequently be generalized and represented on a smaller scale.

The currently (2020) optimal procedure is considered on-screen digitizing. It is a combination of scanning and digitizing. For this procedure, the map as a whole is scanned first and displayed and enlarged on a monitor screen. The lines to be scanned can be highlighted, and the operators will then follow/trace and enter with the cursor the relevant lines or point locations on the screen.

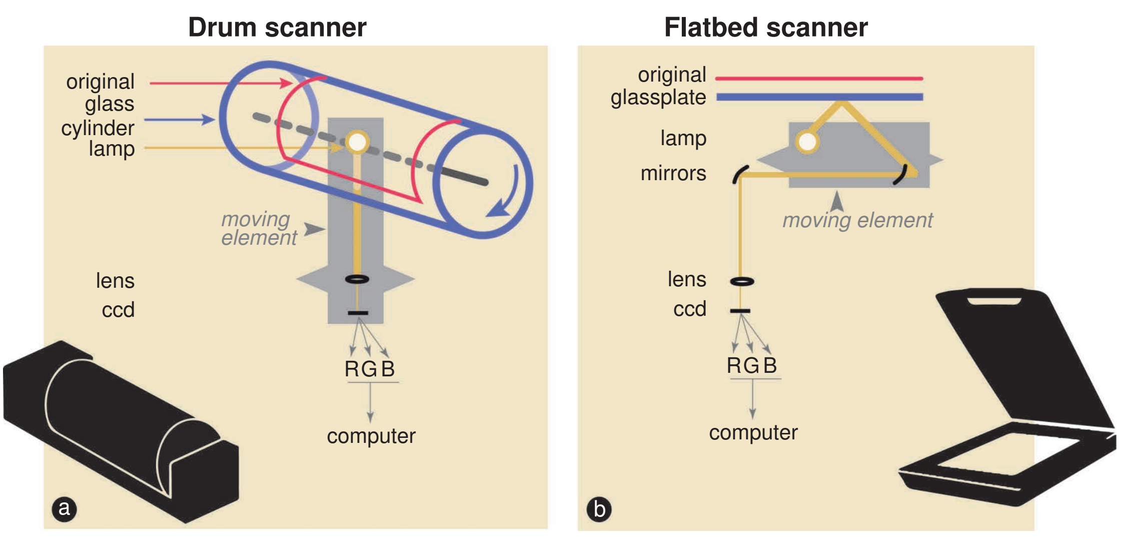


FIGURE 2.8 Working principle of a scanner: (a) drum scanner; (b) flatbed scanner

FIGURE 2.9 On-screen or heads-up digitizing

The parts of the lines entered will be displayed in a different colour; the image can switch between that of the scanned map, partly digitized, and that of the digitized information only, allowing the operator to better check the consistency of work done (see Figure 2.9). This approach is also used by volunteers who want to update map data for humanitarian purposes after disasters such as hurricanes, flooding or earthquakes (in the so-called Missing Maps projects). The volunteers use recent satellite imagery to, for instance, indicate where the damage is most severe, and the new maps, often part of the OpenStreetMap, can be used to restore the basic infrastructure (see also Section 6.6.2).

2.5 WORKING WITH DIGITAL DATA

2.5.1 Modelling the World

Producing today's maps without digital cartographic models is almost unthinkable. As explained in Chapter 1, digital cartographic models are derived from digital landscape models (Figure 1.4). It is the content and structure of these models which determine the possibilities for querying the data and for defining a digital cartographic model needed to draw maps as required. The content of the digital landscape model itself is defined by selections from reality. Obviously, reality as experienced outside cannot be incorporated in a model as a whole. Selectivity is necessary to keep the model workable. In the framework of an application, one will try to process as many aspects of phenomena as possible. The model's complexity depends on the nature of the application and the intended manipulation in the GIS database or the maps required.

How an application influences the contents of a digital landscape model can be illustrated when considering the road concept. To an environmentalist and a traffic manager, a road seems to be the same object. However, the two viewpoints may differ considerably. An environmentalist will look at the road as a barrier to wildlife migration patterns. From this perspective, he or she wants to know its width, how busy it is, the width of the verge, whether there are any crash barriers, the level of noise pollution, etc. A traffic manager will look at the road from a safety and transport perspective. Questions he or she will have are related to the capacity of the road, the number of accidents, traffic lights, flyovers, etc.

Figure 2.10 depicts the modelling process. The illustration is a broadening of Figure 1.3. The steps in the modelling process correspond with the approach suggested by Peuquet (1984). Next to reality, where geographical objects and their characteristics can be found, she distinguishes a data model, a data structure and a storage structure. For each application, selections are made which adhere to conditions defined in a data model. This data model is a conceptualization of reality without conventions or restrictions regarding its implementation. It contains a defined set of geographical objects and their relationships. The example data model comprises the six districts of Cumbria (Figure 2.10a). The next step in the modelling process is to structure the data. In GIS, this implies the representation of the data model in a vector or raster data structure as illustrated in Figure 2.10b. In a vector data structure, the data are organized according to the objects. Geometric characteristics of the data are represented by sets of coordinates, which, in the map image, are connected by lines (the vectors). Labels link the attributes to the geometry. In a raster data structure, data are organized on the basis of a geospatial address. The geometry is represented by the location of grid cells. The address of the cells links the attributes.

The choice of a suitable data structure should be determined by the application. However, in practice, one is restricted to the data structure implemented in commercial software packages. Still, when judging the data structure, one should at least consider points such as completeness, efficiency, lineage, versatility and functionality. Completeness refers to the possibility of representing all selected data, efficiency refers to data accessibility, lineage refers to the way in which the data were collected, and versatility refers to the possibility of adapting the structure to new circumstances. Functionality refers to the operations that can be executed with

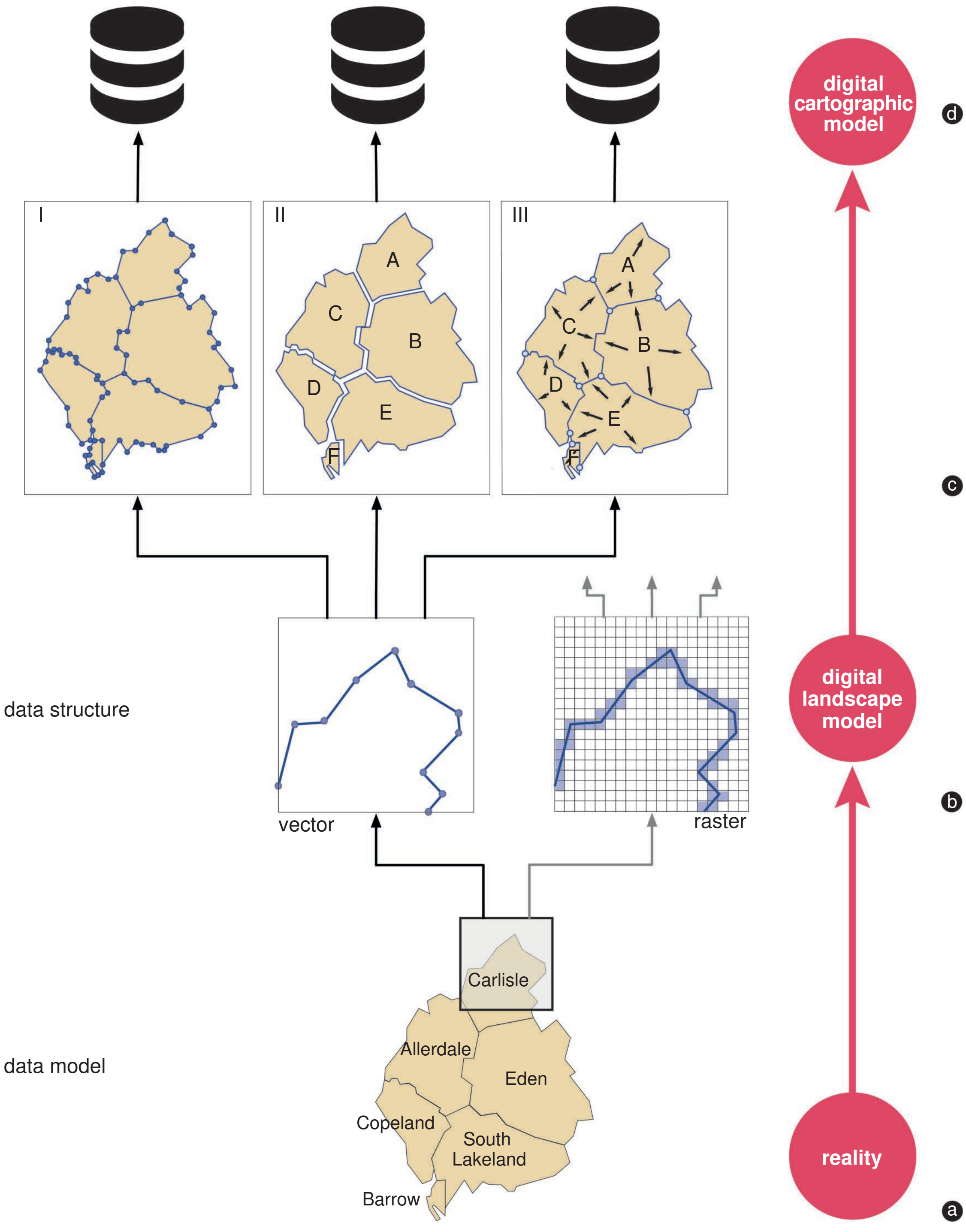


FIGURE 2.10 Organizing geospatial data: (a) selections from reality are based on a data model; (b) the selected data are often structured in a vector or raster format (digital landscape model); (c) the nature of the data structure defines the query level; (d) determination of what maps can be drawn (digital cartographic model)

the data; in other words, it refers to the kind of query that can be answered by the structure. The query level will define the structure's suitability for certain tasks. This is illustrated in Figure 2.10c by three examples of some possible vector data structures in relation to their cartographic use. Map I represents just the individual lines, map II represents the areas as well, and in map III, lines, areas and topology (i.e. mutual relationships) are also known to the system. If one would like to have a simple map with just the outline of the districts, all three data structures will function. If the map to be drawn is based on the request 'draw only the largest and smallest area', the structure represented in map I will fail. However, the structure represented by map II will also fail when the request would be 'draw only those areas bounded by district C'. The type of request will, for raster and vector structures, influence the response times.

The nature of the data to be represented will strongly influence the choice between vector or raster data structures. If one is active in the field of utilities, the vector approach seems an obvious choice, because of the type of objects one is dealing with (e.g. pipes, networks). Whenever the organization depends on remote sensing data, a raster data structure is advisable, because it is suitable for both interpreted and uninterpreted data. It should also be realized that a vector data structure is only suitable for data that have been interpreted fully.

2.5.2 Vector Approach

The vector structure is one of the oldest structures in use. This is partly because the vector approach is close to the traditional cartographic drafting techniques. Look at an analogue map and it is possible to 'see' lines constructed from nodes and arcs. Another reason is the limited computer technology available at that time. The small computer memories were unable to deal with the vast amount of data involved with raster structures, while vector data structures are relatively small. The basic unit of the vector data structure is the geographical object. Several kinds of vector data structure exist. For an extensive elaboration on these structures, see Laurini and Thompson (1992). Figure 2.11 illustrates two of these structures.

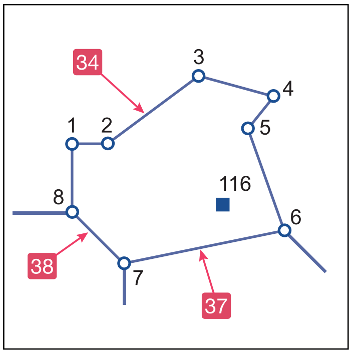
A non-topological or spaghetti data structure is the simplest type of vector structure. All objects are defined as single items. As can be seen in Figure 2.11a, the line between points 6 and 7 is defined by a set of coordinates ($x6$, $y6$) and ($x7$, $y7$) and is labelled 37. However, the data structure does not include any reference to other objects, such as lines 38 and 34, which are connected to line 37. No reference is made to the areas bounded by line 37. It is similar to map I in Figure 2.10c, and it does not refer to any geospatial relationships between the objects defined in the structure. It is unable to answer questions that are not related to drawing its content. Checking its consistency can only be done visually. This approach was introduced at the beginning of the 1970s.

More advanced are those vector data structures that contain topology. Topology defines the mutual relations between geospatial objects and can be used to check consistency among point, line and area objects, or help in finding answers to more complex queries. Topology can also be described as the highest level of generalization possible. Graphically, this can result in different images of the same area, as can be seen in Figure 2.11b. Area C in Figure 2.11b I has its 'natural' boundaries. In II and III, these boundaries have been strongly generalized, but relations between area C and its neighbours are still valid. One of the first vector data structures that included topology was Dual Independent Map Encoding (DIME). This DIME system was developed by the US Bureau of the Census to deal with census data. An important difference to the spaghetti structure is the incorporation of geospatial relationships between the objects registered by the structure. In the 1980s, the Bureau of the Census replaced DIME with the more advanced TIGER (Topologically Integrated Geographic Encoding and Referencing) system.

More complex and advanced is the georelational data structure. Similar to topology, it includes links to a database system containing attribute data. This results in an efficient and flexible structure, and can be found in many of today's GISs, among them ArcGIS. Figure 2.11c shows the principles of this approach. The basic unit is the line segment. From each segment, the beginning and end points (the 'from' node and the 'to' node, respectively) are known. Those two points give direction to the segment and make it possible to define a left and a right area. In Figure 2.11c, the line labelled 17 has point 8 as a begin node and point 6 as an end node. The left and right areas are OUT and A, respectively. When applicable, the number of points between both nodes is registered. The segment labels can be used as pointers to refer to other tables with information on points, areas or attributes. They function as a link to the database as well. The georelational data structure allows for flexible search operations, area aggregation, linking attribute data and consistency checking. The popular file formats GeoJSON and ESRI's shapefiles are kind of georelational data structure, but not quite.

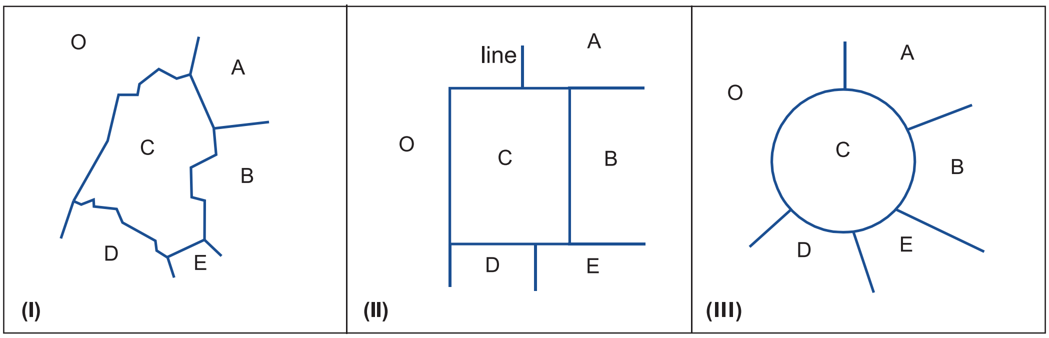
2.5.3 Raster Approach

Since 1990, the use of raster data structures has increased. Although it is difficult to code during input, the speed of the scanner, as well as that of output equipment such as laser and electrostatic a

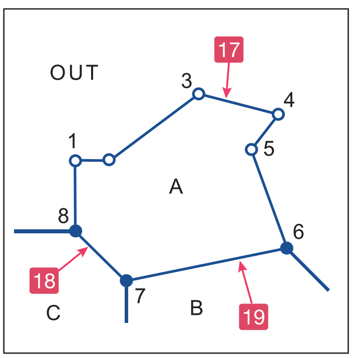


	code	coordinates	attribute
point	116	(x,y)	town
	34	(x_8,y_8) , (x_1,y_1),........, (x_6,y_6)	boundary
line	37	(x_6,y_6) , (x_7,y_7)	boundary
	38	(x_7,y_7) , (x_8,y_8)	boundary

b



c



polygon	area	perimeter	attribute-1	attr.
A	783	1621	name	

line	length	from node	to node	polygon left	polygon right	attribute-1	attr.
17	1032	8	6	OUT	A	external boundary	
18	211	8	7	A	C	external boundary	
19	378	6	7	B	A	external boundary	

points	coordinates	attribute-1	attr.
1	x , y	....	....
2	x , y	....	....
...	...	....	....
...	...	....	....
8	x , y	....	....

FIGURE 2.11 Vector data structures: (a) a spaghetti structure; (b) topology; (c) a georelational structure

plotters, offers advantages. Physical storage is simple, and stored data are easily accessible. A wealth of processing techniques is available from the remote sensing and image processing disciplines. In the GIS environment, this has led to the development of raster modelling techniques (Tomlin, 1990).

In the raster approach, geospatial units function as basic reference units, instead of as geographical objects (as is the case with the vector data structure). Squares are most common (Samet, 1989) since manipulations with squares are easily performed by the hardware (e.g. pixel on a screen or paper). Figure 2.12a shows how a geographical feature is registered in a simple raster structure. The grey squares define Cumbria, and the white squares the area outside Cumbria. Introducing more grey values allows for the registration of more different objects (for instance, different grey values for each of Cumbria's districts). However, it is not possible to register more than one attribute for each square. If there is a need to have more attributes, which occurs especially in a GIS environment, one has to store more raster layers. The size of the squares (the resolution) will define how well a raster structure can represent geographical reality. A small size will give a better representation, but will result in a very large data set. Sometimes the data source defines the resolution. An example is the resolution of the satellite's scanner, which sets conditions for data collected by remote sensing techniques.

Another advantage of the square above a triangle or a hexagon is that it can be split into subunits of the same shape and orientation (Figure 2.12b). The quadtree data structure is based on this approach and is used in GIS packages. Considering the quadtree, the whole map is seen as a square and is subdivided into four smaller squares (see A, B, C and D in Figure 2.12c). Splitting the squares continues until each single square has a homogeneous content. In practice, this goes down to six or seven levels. The example in Figure 2.12c is four levels deep. The tree below the map image shows how it is done. Each solid circle represents Cumbria, and each open circle the area outside Cumbria. When no branches leave a circle, it is considered homogeneous (black or white). Several variations on this basic approach exist.

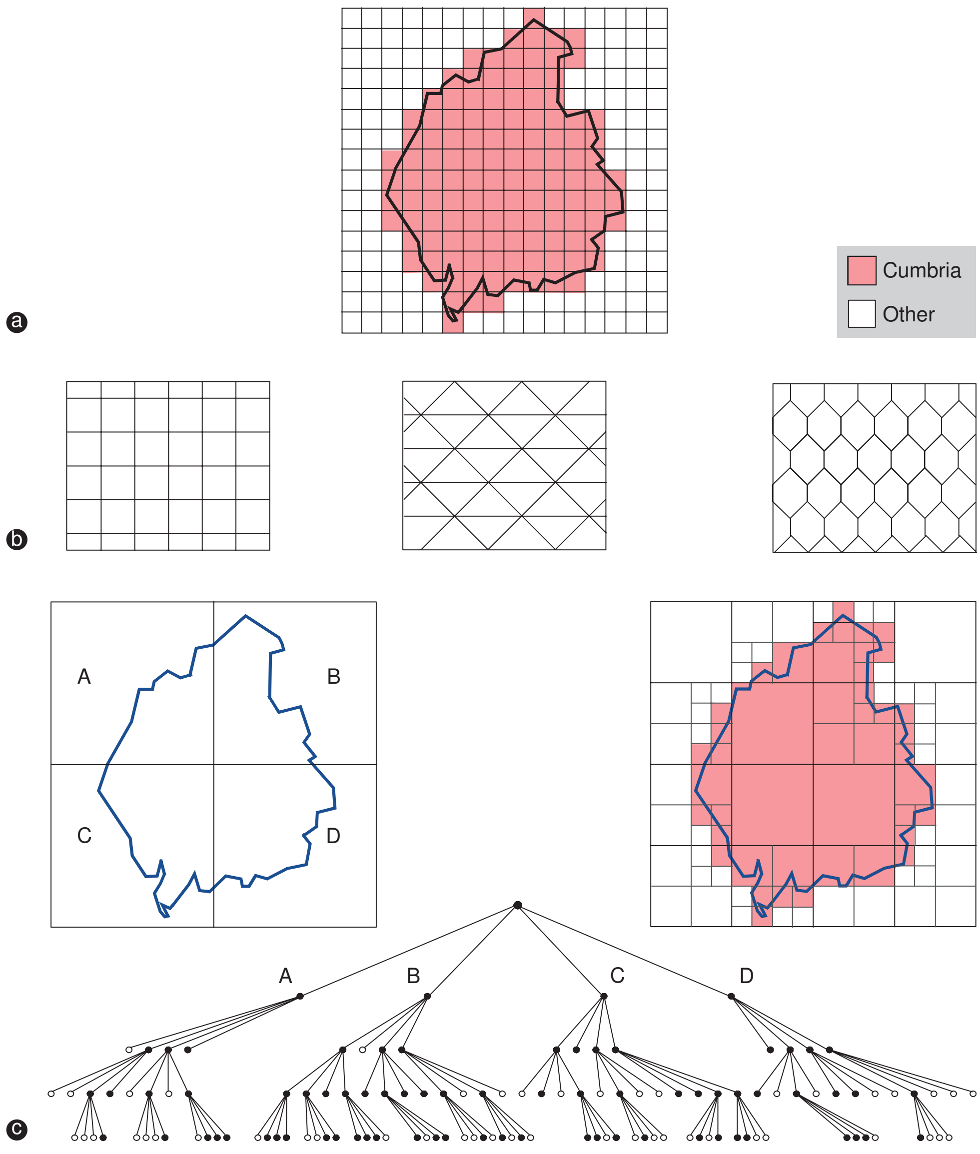
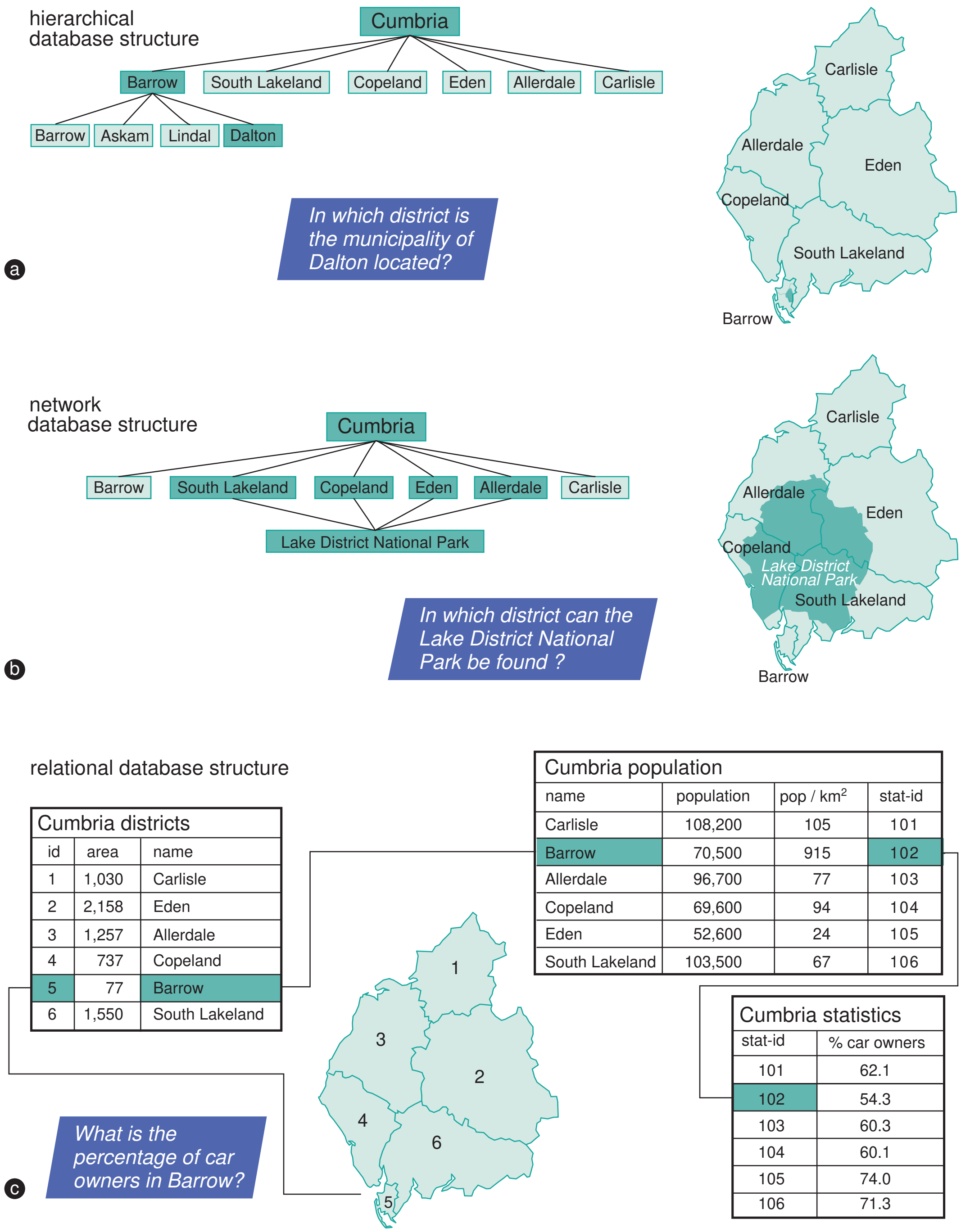


FIGURE 2.12 Raster data structures: (a) normal; (b) basic raster types; (c) quadtree

2.5.4 Hybrid Use of the Database

Both raster and vector data structures are used via database management systems. A database management system often has one of the following structures: hierarchical, network or relational (Figure 2.13). The first has a fixed tree-like structure, and questions can only be asked along the tree's branches, as can be seen in Figure 2.13a. Here, Cumbria is divided into districts, each of which is further subdivided into wards. Because of its fixed structure, it has relatively short response times to queries. It is applied mainly in those information systems with management tasks. Here, the nature of the questions is known beforehand.



Cumbria districts

id	area	name
1	1,030	Carlisle
2	2,158	Eden
3	1,257	Allerdale
4	737	Copeland
5	77	Barrow
6	1,550	South Lakeland

Cumbria population

name	population	pop / km²	stat-id
Carlisle	108,200	105	101
Barrow	70,500	915	102
Allerdale	96,700	77	103
Copeland	69,600	94	104
Eden	52,600	24	105
South Lakeland	103,500	67	106

Cumbria statistics

stat-id	% car owners
101	62.1
102	54.3
103	60.3
104	60.1
105	74.0
106	71.3

FIGURE 2.13 Database management systems: (a) a hierarchical structure allows for fixed queries along the tree's branches; (b) a network structure offers greater flexibility and allows a combination of items on the same level; (c) relational structures are most flexible and allow any kind of question on the data stored in related tables

When a more flexible approach is needed during the query process, a network structure can be used. As Figure 2.13b demonstrates, this structure is not limited to queries along branches. Database elements of the same level, like the districts in the figure, can be combined in the query. These databases can be found at utility companies. However, as with hierarchical databases, the type of question is fixed at the moment one defines the database structure.

A relational database structure is even more flexible. It can handle any kind of query and is very useful in an environment with unpredictable and constantly changing queries. However, it is often slower than the other two database structures. Figure 2.13c gives an example. In the GIS database, the geometry of Cumbria's districts is stored in a table. The relational database principle allows one to link this information with any other table when a common variable is available. This figure shows the district's name and its statistical identifier. The links between the different tables allow one to ask complex questions.

2.6 CONTROL AND ACCURACY

The advent of GIS has accelerated and simplified the process of information extraction and communication. Combining or even integrating various data sets was made possible on a large scale. The ease with which operations can be effected provides a danger as well, as technical possibilities will also allow for irrelevant or inconsistent data integration. On the other hand, the present storage potential in the cloud (storing and accessing programs and data in the Internet instead of in one's computer) will allow one to store the original data and not the derived or aggregated data. Figure 2.14 provides an example of the difference between using original and aggregated data. Here, population distribution data (visualized as a dot map, showing locations of specific numbers of inhabitants) and population density data (visualized as a choropleth map, showing densities for enumeration areas) are both used as a starting point for an analysis of the average distance the inhabitants of a region have to walk in order to reach specific municipal facilities. Therefore, both data sets are combined with

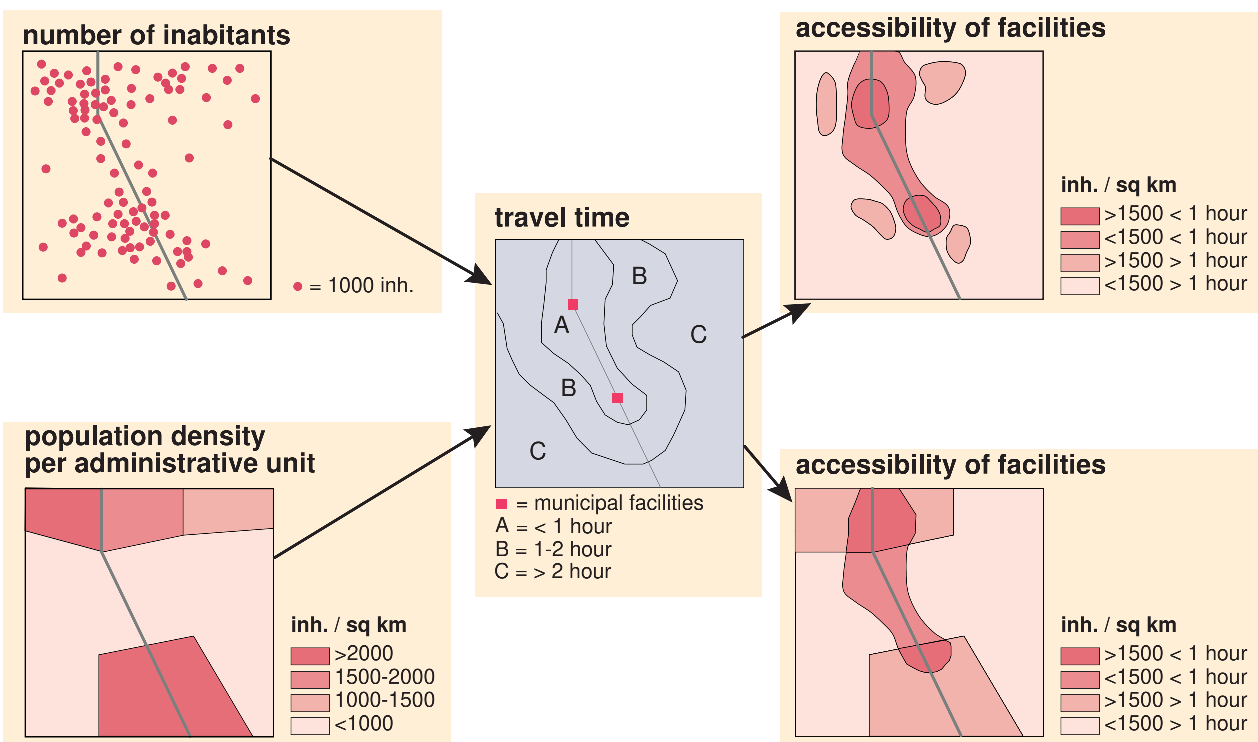


FIGURE 2.14 Results of data integration. If the aggregated data set is used (bottom), the densities of the population able to reach the facilities within a specific time will be visualized differently as compared with the map based on distribution data (top) (after Hootsmans and Van der Wel, 1992)