Lecture Notes in Geoinformation and Cartography

LNG&C

Temenoujka Bandrova
Milan Konecny
Sisi Zlatanova *Editors*

Thematic Cartography for the Society

Springer

Lecture Notes in Geoinformation and Cartography

For further volumes:
http://www.springer.com/series/7418

Temenoujka Bandrova · Milan Konecny
Sisi Zlatanova
Editors

Thematic Cartography for the Society

Springer

Editors
Temenoujka Bandrova
Department of Photogrammetry
and Cartography
University of Architecture, Civil
Engineering and Geodesy
Sofia, Sofiya
Bulgaria

Milan Konecny
Department of Geography
Masaryk University
Brno
Czech Republic

Sisi Zlatanova
GISt, OTB Research Institute
for the Built Environment
Delft University of Technology
Delft, Zuid-Holland
The Netherlands

ISSN 1863-2246 ISSN 1863-2351 (electronic)
ISBN 978-3-319-08179-3 ISBN 978-3-319-08180-9 (eBook)
DOI 10.1007/978-3-319-08180-9
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014941505

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Thematic Cartography for the Society

Introduction

Thematic cartography is one of the most communicative branches of science. It is a large part of cartography focusing on natural phenomena, and social, political, and economic issues, combining visualisation and exploration methods, and targeting and supporting different groups of users. It opens opportunities for the geoinformation society to show its irreplaceable and unique role as the backbone of the information/knowledge based society. It is a tool for decision making and is one of the most suitable ways to show the results of geospatial analysis in a readable visualisation manner. Thematic cartography can support many different disciplines by presenting the knowledge that is obtained from the processing of geoinformation data. Such knowledge transfer is made through maps or cartographical products which gather, manipulate, analyse and visualise information extracted from geodata. Thematic cartography is not simply mapping, but a process which provides users with greater knowledge, in the form of cartographic products: 2D and 3D maps and models of the reality, animations, interactive and web-accessed maps.

In the last decade a broad discussion was initiated about the importance of the user and the need to move from product-oriented to user-oriented cartography. The advances in computer science, computer graphics, machine-human interaction, web and mobile technology have convincingly shown that digital application can be tailored to the needs of the user. Thematic cartography is one of the spatial presentation tools which has long traditions in this direction. Combining newly available technologies with the strength of cartographic theory and practice will allow fast development of user-oriented products in contracts to the traditional application-oriented products. This book provides evidence of how thematic cartography can be used in various areas of society.

The chapters of this book are organised in five parts as follows: User-Friendly Internet and Web Cartography; User-oriented Map Design and Production; Context-Oriented Cartographic Visualisation; Sensing Technologies and their Integration in Maps; Cartography in Education.

1 User-Friendly Internet and Web Cartography

Thematic maps have been always created for the users and with the help of the users. The internet and the world wide web provide new possibilities to make this connection closer and faster. This platform for publishing opens opportunities for everybody to visualise geospatial information through maps. Very often we find maps and cartographic models published on the internet to be of very low quality, insufficient accuracy and of non-professional creation, with a lack of scale, inappropriate map projection, an incorrect symbol system or legend, and content that is difficult to understand. Cartographers need to address these problems and provide users and map-makers with information about map creation and publication for the internet. The validation of information represented on maps originating from volunteers is also very important. The capacity of the internet for publishing is so great that it is sometimes difficult for users to estimate the quality of information, and to choose legible and accurate cartographic information. Cartography has the goal therfeore to provide a user-friendly internet is a platform for developing the most appropriate and useful maps for everybody, everywhere and every time.

The new web-based cartography has the potential to provide society and users with an integrated platform for easy and understandable analyses, spatial data sharing and visualisation. Web cartography can be seen as a part of a spatial decision support system. Mildorf et al. show how an enormous quantity of data coming from different sources could be harmonised, and the result of the processed data offered to the user through thematic maps and web-mapping. For this aim open source tools are used and integrated in an open data platform.

The internet possibilities of modern thematic cartography are of great importance for extending the scope of traditional cartography and this is well-represented in this book. One of the chapters describes a web-oriented geoinformation system for protection against forest fires. Barakovskiy and Zharikova present fire risk maps support urban planners and disaster managers in the prevention phase. Solyman et al. present a methodology for building an integrated web map solution to access the effect of the sea level rise scenarios on the northern coast of Egypt. Meier et al. propose a method for the implementation of density maps as a useful alternative for visualising information on mobile handheld devices. This can help users explore their environment while on the move. The preparation of such maps illustrates also the fact that different specialists need to work together to successfully fulfill the needs and expectations of society.

2 User-Oriented Map Design and Production

Thematic cartography deals with, and makes its products for a variety of users. Several commissions in the International Cartographic Association (ICA) focus on users: for example the commission on cartography and children, education and training, maps and graphics for blind and partially sighted people, maps and society, mountain cartography, planetary cartography, and others. In its research agenda, the ICA aims to present geographic information to society in a user-friendly and understandable visual and tactile way (www.icaci.org).

The major goal of thematic maps is to create relevance and understanding for users. Cartographers need to present geospatial information, in the most clear, readable and appropriate way, to different groups of users with different cultures, religions, ages, genders, etc. Cartography research has published a substantial number of articles about user oriented map production and design over the years. This book also provides several innovative examples.

Dukaczewski presents an approach for designing simple and complex animated maps for users from different age groups, using an appropriate selection of static and dynamic visual and sound variables.

Al-Ghamdi proposes optimisation of the selection of the number of choropleth map classes which could help the GIS user to make better thematic maps. Cartographers use widely GIS because GIS provides them with techniques which allow easy and fast compilation of thematic maps. An example of such work is presented by Kurowska et al., where the universal principles for creating thematic maps supporting the planning process are developed.

Over the last 50-60 years GIS has become one of the most important tools for cartographers in finding solutions to different tasks in society, administration and government. Bartoněk et al. present and analyse the possibilities of optimising the sub-processes and contexts required to determine terrain surface types above gas pipelines in the Czech Republic. They found that in cartographic work, 54% of the time taken to complete a task is still spent on manual work, 23% of the time is fully automated in data preparation and another 23% in data processing.

Rotanova and Lovtskaya discuss the creation of cartographic and thematic databases using GIS, and making a contribution to the SDI of the Ob basin system in Russia.

The usefulness of thematic cartography is demonstrated in the study of Perez-Gomez and Ibanez. They show that thematic maps, 3D models, and zonal statistical analysis are useful alternative methods and techniques when dealing with geodiversity.

Tikunov et al. investigate attempts to systematise elements that form the power of the state, and bring them into one formula. They introduce a general geographic size index (GSI) which includes only three components: area, population and economic strength. GSI demonstrates the real picture of government power and illustrates the dynamics of its change. The results are demonstrated through GSI graphs and maps.

Very often a large team of different specialists needs to work together to produce good thematic products that fit the needs of specific users. A good example of this is the compilation of the Academic Atlas of Czech History which constitutes a unique multidisciplinary publication (Janata et al.).

User-oriented map design and production cannot be performed without the user. Specialists from different domains need to work together with users, and map design and production must involve access to all necessary data with respect to the tasks society needs to perform and the challenges that different groups of society will face.

3 Context-Oriented Cartographic Visualisation

Context-oriented cartography is part of so-called ‘ubiquitous mapping’ which comes from the idea of pervasive computing and pervasive maps. Ubiquitous mapping is one of the biggest challenges in contemporary cartography and geoinformatics, and arose because of the perceived advantages of an ‘information society’. Mapping should be done by everybody, at any time, and everywhere with all possible technological tools.

The main objective of visualisation is to present the salient features of a spatial context in a manner that is easily and quickly assimilated by a human being. The problem lies in the volume of information and the consequential information overload, leading to a user missing vital aspects in their analysis. In a dynamic situation, the problem is compounded by the real-time stream of data that has to be rapidly surveyed by the user for patterns and anomalies without being overwhelmed.

Researchers have discovered and confirmed the cognitive styles of users with different skills, abilities, education and cultural background and different ages, which is especially important in dynamic geovisualisation (e.g. in disaster management) for creating adaptive and context-based map concepts.

Cognitive style or ‘thinking style’ is a term used in cognitive psychology to describe the way individuals think, perceive and remember information, or their preferred approach to using such information to solve problems. The two main streams of cognitive style are holistic or analytic, and information about both could give practical input to cartographers. Some people find maps preferable, others prefer orthoimages. This preference can play an important part in cartography and map using, especially in disaster situations. Several chapters in this book are devoted to this topic.

Yilmaz et al. investigate different means of geovisualisation in a virtual environment, using human perception. Zheng and Zheng’s research describes Open StreetMap data for Chinese territory. They investigated which parts of the biggest cities have the most complete information and which areas of China have less detailed information.

Hu et al. developed a dynamic map environment with three directions: integration of the virtual geographic environment with volunteered knowledge construction, the connection of the dynamic environment to everyday life, and the provision of a comprehensive social experience of daily communication.

Todd developed a common visual symbolic language to represent eco-regions. This will help to represent the biophysical complexity of regions in a simplified form and will also help users making land use decisions.

A simplification of 3D Maps with multiple LODs of buildings is shown in the research of Noskov and Doytsher. They propose a method for compiling a three dimensional scene with automatic LOD generation of objects. This research contributes to improving visualisation performance by reducing the computing time and the required computer resources.

One of the most important subjects of thematic cartography is mapping for the needs of early warning and disaster management. Many authors direct their attention here. A good example of cooperation between scientists from different countries (Bulgaria and Germany) and comparative analyses of different territories (Bulgaria and USA) is the chapter written by Boyanova et al. Such common efforts will bring cartography to the level where it can be used for fast decision making. Cartographic products and analyses are presented for use towards this aim.

4 Sensing Technologies and their Integration with Maps

The integration of sensing technologies and thematic cartography expands the abilities of research and development to support decision makers and the geoinformation-oriented society. Such integration will benefit from being process-oriented, and from mapping of the required territories that uses better resolution maps and more attributes in databases. The economic value of thematic maps for end-users can thus be found. This will come from more detailed spatial information, higher accuracy and additional content which can be extracted from the range of cartographic products made on the basis of this process of integration. Cartographers need to know the end-user requirements so as to present individual objects and phenomena on their maps.

Rocchini et al. describe the integration of sensing technologies with maps and estimate potential hotspots of diversity, allowing effective management and conservation of the landscape. Sensing technologies in combination with cartography have another common target: to describe the results of climate change. Hidayati et al. analyse image transformation and land use cover according to temperature trends shown by Landsat imagery, and find a 0.99ºC rise in average temperature in twenty years in an Indonesian city.

Another chapter considers environmental data visualisation. Lienert describes the linking of real-time sensor data with spatial data infrastructures for web-based visualisation. This responds to the growing expectations of society for spatial data access from sensors. It is achieved through the creation of mobile device-based maps on a public web-based platform.

One of the most important and interesting topics of thematic cartography is the creation of environmental noise maps. This topic is well covered by Duda who uses crowd-sourcing for the creation of such maps. Such approach leads to many questions about the quality of data and its validation but Duda believes that crowd-sourcing a promising low-cost and quick method, which will further develop.

5 Cartography in Education

One of the greatest goals of cartography is in producing maps for education. From a young age students deal with and use maps in a variety of courses in their study. The use of cartography in education processes continues into university and the various courses for specialists from science, government and regional institutions. Cartographers try to use the latest achievements of technology and give their users modern materials using the internet, mobile devices, and animated and multidimensional representations.

Maps for education are becoming more important than ever. Students are familiar with computers and social media. Using games for education is already a trend in many countries. These developments have also influenced the types and content of maps.

Rodionova et al. show that thematic cartography gives users a great deal of new information about the surfaces of celestial bodies, and they describe in detail the mapping of extra-terrestrial objects. Their research is also connected to education and the university teaching of cartography.

School cartography is considered by Reyes Nunez. His research is about the use of cartograms as one of the newest methods of representation in thematic cartography, and in maps compiled for students.

Ormeling describes different types of distortion in cartography: hill-shading, classification, generalisation, and map projections. He recommends "more warnings attached to maps regarding these drawbacks, and [that] the unsuitability of the use of these maps for specific purposes, such as making measurements, should be clearly stated. It should be easy to attach such warnings to the maps, especially to their digital files". The same warnings should be addressed to mapmakers who are not cartographers.

As the chapters in this book illustrate, cartography is a discipline which has always taken into consideration the needs and the requirements of users. This unique role should be further strengthened by close cooperation with specialists from other disciplines so as to be able to provide the best products for society.

We hope that this volume will be useful reading for researchers, practitioners and students. We are grateful to the contributors for their commitment and hard work. We greatly appreciate the support of all of the members of the International Organising Committee of the 5th Jubilee International Conference on Cartography and GIS and are thankful for their time and valuable comments, which have contributed to the high quality of this volume.

Our sincere thanks go to Stephan Angsüsser (China), David Fraser (Australia), Philippe De Maeyer (Belgium), Milan Konečný (Czech Republic), Horst Kremers

(Germany), Petr Kubíček (Czech Republic), Miljenko Lapaine (Croatia), Eugene Levin (USA), Lyubka Pashova (Bulgaria), Necla Ulugtekin (Turkey), Laszlo Zentai (Hungary), Sisi Zlatanova (The Netherlands) for their reviews and to Stefan Bonchev (Bulgaria) for his technical assistance.

This book is the result of a collaborative effort of 71 researchers from Bulgaria, China, Czech Republic, Egypt, France, Germany, Hong Kong, Hungary, India, Indonesia, Israel, Italy, Netherlands, Poland, Russia, Saudi Arabia, Spain, Switzerland, Turkey, Ukraine, and the USA.

Temenoujka Bandrova, Milan Konecny, Sisi Zlatanova

Part I
User-Friendly Internet and Web Cartography

Open Data Platform for Data Integration, Visualisation and Map Design

Tomas Mildorf, Jan Jezek, Otakar Cerba, Christian Malewski, Simon Templer, Michal Sredl, Karel Charvat[1]

Abstract The current trend in the EU is to open access to public sector information which is provided either for free or for marginal cost, and reuse it in various applications. Information technologies enable people to access, process and analyse spatial data from various sources, help to design on-demand maps and provide information for decision makers. However, the provision of data varies across different authorities, and combining heterogeneous data is not an easy task. We present an Open Data Platform that enables people to integrate, harmonise and visualise spatial planning and other data. The platform connects to the approach of real cartography and aims to enable non-cartographers to correctly design maps and gain new information in a user-friendly way based on modern technologies and robust data storage. This chapter mainly tackles the issues of heterogeneous data integration, harmonisation and visualisation. Ongoing research aims to explore new methods of data reuse and cartographic visualisation, following the trends of modern cartography.

1 Introduction

Spatial data is a key part of knowledge about the relationships between phenomena on the Earth. Spatial data links various kinds of information and provides the means for data interpretation and visualisation. The availability of spatial data is

---

T. Mildorf, J. Jezek, O. Cerba
University of West Bohemia, Faculty of Applied Sciences
Univerzitní 8, 306 14 Plzeň, Czech Republic

C. Malewski, S. Templer
Fraunhofer Institute for Computer Graphics Research IGD
TU Darmstadt, Interactive Graphics Systems Group
64283 Darmstadt, Germany

M. Sredl, K. Charvat
Help Service Remote Sensing
256 01 Benešov, Czech Republic

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_1,

thanks to significant improvements in initiatives such as INSPIRE[2] or Open Data[3], and access to actual and clearly licensed data is becoming easier. Spatial data is available not only for viewing purposes but also for reuse, and can be published as machine readable, which opens the field for innovative services based on linked data[4].

The current trend in the EU is to open access to public administration information. Information is provided either for free or for marginal costs. In addition to EU legislation supporting this process, such as the EU Directive on the Re-Use of Public Sector Information (PSI Directive), there are other European initiatives, including, for example, the European Interoperability Framework (EIF). All related legislation and initiatives aim to provide information in an interoperable way that is suitable for reuse (Mildorf et al. 2013). As identified in recent studies, for example by Koski (2011), the release of public sector information for reuse can be of considerable benefit to economic growth.

The information value of data increases when different data sources are connected or linked together. New information can be then derived through data mining. The ability to effectively integrate, interpret and visualise many different datasets is one of the major challenges for modern cartography and geoinformatics.

There are several platforms that focus on data access and interpretation, such as the World Bank[5] or the World Factbook[6]. Such platforms focus on the visualisation and comparison of data from particular countries or groups of countries. We present a platform that specialises in spatial planning data used for, or resulting from, spatial planning activities and their reuse.

Spatial planning "gives geographical expression to the economic, social, cultural and ecological policies of society. It is at the same time a scientific discipline, an administrative technique and a policy developed as an interdisciplinary and comprehensive approach directed towards a balanced regional development and the physical organisation of space according to an overall strategy." (Council of Europe 1983). Spatial planning data, as understood by the authors, includes mainly land use data, as defined by the INSPIRE Directive: "the territory characterised according to its current and future planned functional dimension or socio-economic purpose (e.g. residential, industrial, commercial, agricultural, forestry, recreational)." (European Commission 2010). As well as land use data, spatial planning data encompasses statistical data, hydrography, flood areas, protected sites, transport networks, cadastral parcels and other data used for spatial planning activities.

---

[2] http://inspire.jrc.ec.europa.eu/

[3] http://opendefinition.org/

[4] http://linkeddata.org/

[5] http://data.worldbank.org/

[6] https://www.cia.gov/library/publications/the-world-factbook/

In this chapter, the concept of a platform that aims for easy data integration into a large database of open data is introduced. Tools and applications for data visualisation can be built, based on this database. The platform was developed within the Plan4business project[7]. The platform tackles the all-important areas of data integration, harmonisation and map design to map accessibility through user friendly searches. The emphasis is on a user-oriented approach and interoperability through the Open Geospatial Consortium (OGC) and other standards.

In Section 1 the trends of modern cartography are presented. Section 2 introduces the general architecture of the open data Platform. Section 3 describes the harmonisation part of the platform and its features. Section 4 describes the tools for publication and map design. Conclusions are presented in the last section.

2 Trends in Modern Cartography

The traditional concept of cartography mainly tackles the issue of spatial data visualisation. Modern cartography extends the traditional concept by focusing on the communication and exploration of spatial information. This is documented in the cartographic cube (MacEachren 1994), which shows that this shift is connected to interactivity and public involvement.

There are many trends and directions in contemporary cartography that are connected with the research introduced in this chapter. First of all, it is necessary to mention the web or internet cartography that has brought a change in the role of maps or the way maps are used (Kraak & Brown 2000, Peterson 2003).
A complete overview of contemporary cartographic trends is published in the ICA Research Agenda (Virrantaus et al. 2009). The solution proposed in the Plan4business project is related to many fundamental keywords and terms in the ICA Research Agenda, such as map production, geovisualisation, visual analytics or usability.

The research and final products also conform with the "Research areas for next years" and "Cartography 2020" concepts published for the GeoCart 2012 conference (Auckland, New Zealand) by Georg Gartner (the President of the International Cartographic Association (ICA)).

Another important direction of map making is related to data journalism (Gray et al. 2012). This approach is based on the development of maps and data visualisation products that reduce cartographic rules and principles, for maps that are constructed by non-cartographers. Data journalism emphasises a simple method of map building and the aesthetics of final products.

The proposed platform connects both the above-mentioned approaches - real cartography (1), and an effort to enable non-cartographers to make accurate maps

[7]http://www.plan4business.eu/

and gain new information in a very simple and user-friendly way based on modern technologies and robust data storage (2). The development of the platform is on-going and the authors are attempting to find a new, unusual and attractive means of map design which is understandable by a wider audience (Konečný 2013).

The proposed solution also tries to eliminate the gaps identified by Kraak and Ormeling (2010) between technological opportunities (e.g. map portals, large databases or advanced analysis tools) and cartography (above all cartographic theory).These gaps could be reduced by using numerous cartographic techniques and methods of contemporary thematic cartography.

The authors draw on the "Czechoslovakian cartographic school", which distinguishes many types and sub-types of well-known cartographic techniques (such as choropleth or diagram maps) published by Pravda (2006) or Voženílek and Kaňok (2011).

3 Open Data Platform

The key issue for map design is to begin with appropriate spatial data. The current situation, which is underpinned by several initiatives including INSPIRE, GEOSS and Copernicus, shows the heterogeneity of spatial data from various sources and also the difficulties in integrating this data for further reuse. Such reuse includes making derivatives from the data and using them for analyses or as a base layer for other applications. These general problems led the Plan4business team to design and develop a platform capable of the integration of data from scattered sources, and the provision of this data for reuse. The platform mainly integrates open data available at European, national, regional and local levels.

The main incentive for the platform design was the possibility of integrating spatial planning data from various sources, so as to store it in a harmonised manner and to provide it in an interoperable way for further applications such as map design and spatial analyses.

The platform is composed of three layers including the integration engine, storage engine and the analysis engine (see Fig. 1.1). The integration and storage engines use the INSPIRE Data Specifications on Land Use (INSPIRE TWG Land Use 2013) in order to achieve data interoperability for reuse and to enable cross border analyses and the design of maps in a common style. These engines are described in detail in the next chapters.

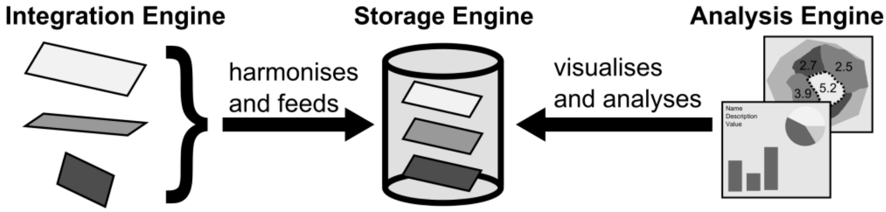


Fig. 1.1 Composition of the three core engines; integration engine for data harmonisation, storage engine for data storage and provision, and the analysis engine to visualise and analyse harmonised data.

4 Data Integration and Storage

The integration engine and parts of the storage engine create harmonisation in the Open Data Platform. The integration engine's task is to provide tools to effectively transform data from any schema into a pre-defined target schema. The storage engine allows for the management and provision of this harmonised data. The latter also manages non-harmonised data and application-specific data (Gregor et al. 2014).

The current version of the integration engine focusses on the harmonisation of heterogeneous land use data in European countries. Land use data might involve either the current status of a particular region's land use or a land use plan that determines the future use of particular regions, as defined in the INSPIRE Specification on Land Use (INSPIRE TWG Land Use 2013). The latter was chosen as a functional basis for the land use data models in Plan4business. Since the results of the INSPIRE initiative will prospectively have legal character for public authorities, the Plan4business data model is designed with compatibility to that future standard in mind. Moreover, harmonisation is adjustable with medium effort in order to address other themes in the 34 that comprise the INSPIRE directive.

Data harmonisation is achieved through transforming spatial data sets based on a set of schema mapping instructions. These mapping instructions may include, for example, reclassifications of land use nomenclatures, spatial coordinate transformation and the assignment of object types from the source data to a target schema. The integration engine has been realised as a web service based on the *Humboldt Alignment Editor* (HALE) software stack[8]. HALE provides the functionality to perform interactive mapping of geospatial schemata. HALE's user interface was adopted in order to allow the mapping process to be performed online. A step-by-step guide will therefore lead the user through the mapping process and help to map the source entity types to the target schema.

Data transformation is performed in three steps. The data for transformation is first uploaded to the integration engine. The upload process includes determina-

[8]http://www.esdi-community.eu/projects/hale

tion of the spatial reference system. A pre-configured HALE mapping project is loaded into HALE in a second step, where a mapping instruction is defined manually for each source schema's entity and attribute type with a corresponding type in the target schema. This mapping is based on the user's domain knowledge. In a third step the final mapping project is uploaded to the integration engine and applied to the source data set, resulting in a transformed data set that is INSPIRE compliant.

One of the core aspects of the land use domain, in addition to spatial coordinate transformation and the retyping of entities and attributes, is the re-classification of land use categories (Fig. 1.2). Re-classification can be realised through mapping-tables connecting a source classification to one or more classification categories of the Hierarchical INSPIRE Land Use Classification System (HILUCS). This process is not straightforward, since mapping is not necessarily unambiguous and domain experts might map classification values differently. Loss of information may also occur when a category has to be mapped to a more general one in the HILUCS system. To overcome these issues our approach is to develop collaborative functionality for the re-classification process based on social network functionality, where domain experts can discuss and rate available re-classifications, and develop mapping strategies for controversial cases.

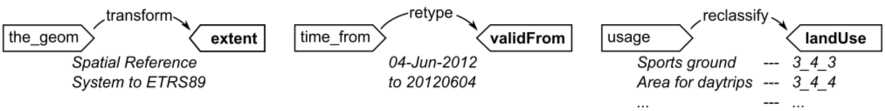


Fig. 1.2 Exemplary mappings for spatial, temporal and thematic attributes. The bold attribute names belong to the target schema INSPIRE Data Specification on Land Use.

The data represented in the target schema is analysed and validated before it is sent to the storage engine and managed in a relational database model similar to the target schema. At the end of the data processing pipeline, the data is either accessible as INSPIRE compliant files (e.g. Geography Markup Language[9]) or via SQL.

As well as the core functionality of managing harmonised land use data in a relational data base, the storage engine provides additional data and functionality. On the one hand, it offers data from other themes, such as statistical data, or, for example, flood zones that can be combined with the land use data in order to run customised queries and create map overlays. On the other hand, the storage engine also includes a graph database. Selected data sets can be made accessible and searchable via URIs and SPARQL queries. This raises the system to the level of linked data.

[9]http://www.opengeospatial.org/standards/gml

5 Map Design and Publication

Harmonisation in the Open Data Platform currently supports land use data as described in the previous chapter. A large pool of land use data, for both current and future land use (urban plans), is being created, and serves as the basis for map creation, spatial analyses and data reuse.

As well as land use data, the Open Data Platform contains other thematic data including:

- statistical data – subsets of data provided online by EUROSTAT and selected national statistical offices,
- land cover (land use) data – CORINE Land Cover and Urban Atlas distributed by the European Environmental Agency,
- cadastral data – currently limited to data from the Czech Republic,
- transport data – data extracted from the Open Street Map,
- flood data – 5-year, 20-year, 100-year and the biggest flood zones in the entire territory of the Czech Republic.

Common application schemas for these data are being designed in order to make reuse of this data possible in a cross-border context This should help, as in the case of land use data, to simplify the processes of data integration and analysis. Once data is integrated and stored in the database in an interoperable way, it can be used for numerous applications, such as map design and map publication. For this reason, the Plan4business team has implemented a web based tool (Layer Manager) for data management including access control, cartographic map design and web publication.

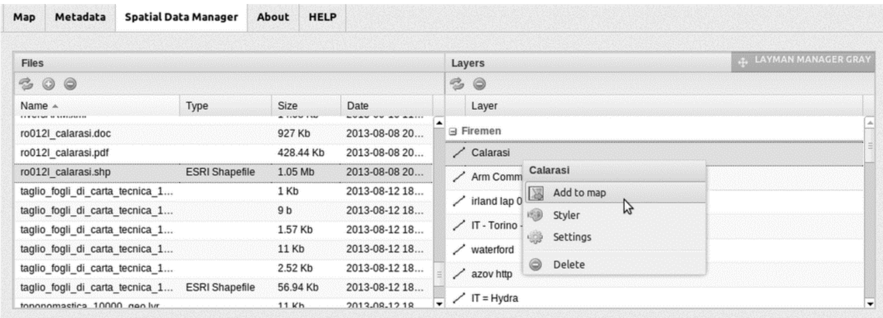


Fig. 1.2 Layer Manager: uploaded spatial data on the left side and published map layers on the right side.

The Layer Manager (Figure 1.3) is an integral part of the Open Data Platform and is responsible for data management in the context of the GeoServer installation. The integration engine makes use of the Layer Manager API (Application Programming Interface) to publish integrated land use datasets through GeoServer. Data other than land use can be directly uploaded and published through the Layer Manager without any interaction with the integration engine. Users can also change the style of the map using the Styler tool.

Data that is either integrated using the integration engine or directly uploaded through the Layer Manager can be visualised through a map viewer or OGC web services and used for designing user defined map output.

6 Conclusions

The platform presented offers a means for data harmonisation, analysis and sharing, in the form of interoperable data, thematic maps and visualisations. The stress is on automation of the processes and intuitive interaction with the user.

The diversity of data coming from various sources is enormous. In order to increase the value of spatial data, integration and links with other data is necessary. The platform aims at automation of data integration. Once mapping between the original and the target application schema is defined, the integration can be performed automatically, which is the advantageous key aspect of this platform. As soon as data is harmonised, analyses, visualisations and cross border applications can be built on top of it.

The entire map production process, from data integration to map design, is based on Open Source tools integrated into the Open Data Platform. The sustainability of this platform should be secured by its users, who will provide input for the data pool (e.g. integrate local urban plans) and who will use its other features as output, for such as web mapping, thematic map creation, map design or spatial analyses.

Acknowledgments This chapter was supported by the FP7 project Plan4business: “A service platform for aggregation, processing and analysis of urban and regional planning data” (Grant agreement no. 296282); the European Regional Development Fund (ERDF), project “NTIS – New Technologies for the Information Society”, European Centre of Excellence, CZ.1.05/1.1.00/02.0090; the CIP-ICT-PSP-PB project SDI4Apps (621129); and the EXLIZ project – CZ.1.07/2.3.00/30.0013, which is co-financed by the European Social Fund and the state budget of the Czech Republic.

References

Council of Europe (1983) European regional/spatial planning Charter. http://www.coe.int/t/dgap/localdemocracy/cemat/VersionCharte/Charte_bil.pdf. Accessed 2 May 2014

European Commission (2010) Feature Concept Dictionary. https://inspire-registry.jrc.ec.europa.eu/registers/FCD. Accessed 25 March 25 2013

Gray J, Chambers L, Bounegru L (2012) The Data Journalism Handbook. How Journalists Can Use Data to Improve the News. O'Reilly Media, Sebastopol

Gregor R, Klien E, Malewski C, Mildorf T, Šrédl M, Templer S (2014) D5.2 Integration Engine. Project report, Plan4business consortium. http://www.plan4business.eu/cgi-bin/download.pl?f=208.pdf. Accessed 2 May 2014

INSPIRE TWG Land Use (2013) D2.8.III.4 INSPIRE Data Specification on Land Use. Technical Guidelines. http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_LU_v2.0.pdf. Accessed 2 May 2014

Konečný M (2013) Maintaining good quality maps with new techniques is a major concern. http://www.geospatialworld.net/Interview/ViewInterview.aspx?id=30757. Accessed 13 Jan 2014

Koski H (2011) Does Marginal Cost Pricing of Public Sector Information Spur Firm Growth? Research Institute of the Finnish Economy, Helsinki.

Kraak MJ, Brown A (eds) (2000) Web Cartography: developments and prospects. Taylor & Francis, London

Kraak MJ, Ormeling F (2010) Cartography. Visualization of Spatial Data. Harlow, Pearson

MacEachren A (1994) Visualization in modern cartography: Setting the Agenda. In: MacEachren A, Taylor D R F (eds) Visualization in Modern Cartography. Oxford, Pergamon

Mildorf T, Ježek J, Kepka M, Čerba O, Klien E, Templer S, Charvát K (2013) Plan4business - a service platform for aggregation, processing and analysis of urban and regional planning data. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol XL-4/W1, p41-44

Peterson M (ed) (2003) Maps and the Internet. Elsevier Press, Cambridge

Pravda J (2006) Metódy mapového vyjadrovania, klasifikácia a ukážky: Geographia Slovaca 21 (Map representation methods, classifications and examples). Geografický ústav SAV, Bratislava

Virrantaus K, Fairbairn D, Kraak MJ (2009) ICA Research Agenda on Cartography and GI Science. The Cartographic Journal, vol 46, number 2, May 2009, Maney Publishing, p 63-75

Voženílek V, Kaňok J (2011) Metodytematickékartografie –vizualizaceprostorovýchjevů (Methods of thematic cartography – visualisation of spatial phenomena). UniverzitaPalackeho, Olomouc, p 216

A Web-Oriented Geoinformation System Application for Forest Fire Danger Prediction in Typical Forests of the Ukraine

Nikolay Baranovskiy and Marina Zharikova

Abstract A web-oriented geoinformation system for forest fire danger prediction based on a probabilistic fire danger criteria is described in this chapter. A new method for determining the probabilistic fire danger criteria is described. A new formula for fire danger assessment for the j-th time interval of forest fire season is obtained using the basic principles of probability theory. A definition of probability using frequency of events is used to calculate fire danger. Statistical data for certain forests is used to determine all the multipliers in the formula for fire danger. The system is developed in the Django platform in the programming language Python. The system architecture, based on Django's Model-View-Template, is described. The software package that runs on the server allows a set of parameters describing forest fire danger to be obtained and used for visualisation. A part of forest fire risk map which correspond to certain value of fire danger is depicted. Estimation of fire risk helps to identify the areas most prone to fire ignition, so as to efficiently allocate forest fire fighting resources.

1 Introduction

Forests and forest ecosystems are of key importance for the social, economic and environmental viability and development of each country of the world. Forest fires are considered a major and permanent threat to forests. Forest fires have negative impacts, which last from the combustion period up to decades after, especially for large fires. The problem of forest fire protection is one of the most important and

N. V. Baranovskiy
Department of Theoretical and Industrial Heat Systems Engineering, Institute of Power Engineering, National Research Tomsk Polytechnic University, Tomsk
Russian Federation
e-mail: firedanger@yandex.ru

M.V. Zharikova
Department of Information Technologies, Kherson National Technical University
Kherson, Ukraine
e-mail: marina.jarikova@gmail.ru

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_2,

includes a wide range of actions which isn't limited to reactive measures such as fire extinguishing (Grishin 1981). According to the definition given in Simard (1976) forest fire protection is a set of actions aiming to reduce the negative impact of forest fires on natural resources, ecosystems and the environment, according to the goals of the management organisation. The main tasks in protection against forest fire are forest fire prevention, extinguishing and minimization of the consequences. Forest fire prevention is a major task which is of great practical importance and requires detailed study. Forest fire prevention is impossible without fire danger assessment. Contrary to other natural hazards (earthquakes, storms etc.), forest fires are generally predictable, but this opportunity has not yet been properly utilised. In this connection we propose new probabilistic fire danger criteria. Spatial planning is of primary importance when decision-making for forest fire prevention. Towards this end we have created a web-oriented geoinformation system for forest fire danger prediction based on the probabilistic fire danger criteria. The system was developed on the Django platform in the programming language Python.

2 Background

Decision-making in wildfire prevention is based on the huge amount of knowledge which describes the complicated interdependencies between physicochemical, biological, social and economic processes. These processes are characterised by stochastic behaviour, and spatial and time dependency. Wildfire prevention is a critical domain where incorrect management decisions may have disastrous social, economic and ecological consequences (Khodakov and Zharikova 2012, Kuznetsov and Baranovskiy 2009). A decision-maker has to analyse a huge amount of information when making decisions under conditions of uncertainty. In such cases decision-makers often make their decisions based on previous experience. The method of fire danger assessment proposed in this chapter, is based on the processing of statistical data reflecting this experience. The function of the geoinformation system described in the chapter is also based on this method. The purpose of the chapter is to describe the method of probabilistic fire danger criteria and the geoinformation system based on this method.

3 Probabilistic Fire Danger Criteria

Using the basic principles of probability theory (Nazarov and Terpugov 2006) the author obtained a formula for fire danger assessment for the j-th time interval of the forest fire season (Baranovskiy 2007):

$$P_j = \left[P(A)P(A_j/A)P(FF/A, A_j) + P(L)P(L_j/L)P(FF/L, L_j)\right]P_j(D), \quad (1)$$

where P_j is the probability of a forest fire beginning for the interval j in the controlled forest area; *P(A)* is the probability of a fire source from an anthropogenic cause; $P(A_j/A)$ is the probability of a fire source on a week day *j*; $P_j(FF/A,A_j)$ is the probability of a forest fire starting from anthropogenic causes in the forest area; *P(M)* is the probability of dry thunderstorms in the forest area; $P(L_j/L)$ is the probability of ground lightning discharge; $P_j(FF/L,L_j)$ is the probability of a forest fire beginning from lightning on the assumption of a dry thunderstorm in the forest area; $P_j(D)$ is the probability that the forest fuel layer will be dry; index *j* corresponds to a day of the fire danger season.

The author proposes to use a definition of probability involving frequency of events and using statistical data for certain forests to determine all the multipliers in the formula (1).

The formula contains the following members (Baranovskiy, 2007):

$$P(A_j / A) \approx \frac{N_{FD}}{N_{FW}}, \; P(A) \approx \frac{N_A}{N_{FS}}, \quad (2)$$

$$P_j(FF / A, A_j) \approx \frac{N_{FA}}{N_{FT}}, \quad (3)$$

$$P(L_j / L) \approx \frac{N_{LH}}{N_{LD}}, \quad P(L) \approx \frac{N_L}{N_{FS}}, \quad (4)$$

$$P_j(FF / L, L_j) \approx \frac{N_{FL}}{N_{FT}}, \quad (5)$$

where N_A is the number of days during fire danger season when the anthropogenic load is enough for forest fuel ignition; N_{FA} is the number of fires from anthropogenic load; N_{FT} is the total number of fires; N_L is the number of days when there was lightning (when dry thunderstorms took place); N_{FS} is the total number of days in the fire season; N_{FL} is the number of fires from lightning (when dry thunderstorms took place); N_{FD} is the number of fires on a specific day of the week; N_{FW} is the total number of fires during a week; N_{LH} is the number of passing ground lightning discharges from 00.00; N_{LD} is the total number of passing ground lightning discharges per day.

Obviously, the more cases considered for a certain forest, the more precise formulas (2)-(5) will be. All parameters of the fire danger season (N_A, N_{FA}, N_{FT}, N_L, N_{FS}, N_{FL}, N_{FD}, N_{FW}, N_{LH}, N_{LD}) will therefore have to be registered every year.

Formula (1) contains the multiplier $P_j(D)$. This is the probability of fire danger from meteorological conditions. In early work, this probability was calculated using the time taken for the fuel layer to dry (Grishin and Baranovskij 2003). Implementation of this method throughout the Russian Federation and the Ukraine is difficult, however, because to model the process of forest fuel drying we need information about the initial moisture content of forest fuel. We propose a compromise version in this chapter. It is proposed to calculate the probability of meteorological conditions using the Complex Meteorological Index, which was approved in the Russian State Standard 22.1.09-99 "Monitoring and forecasting of forest fires. General requirements" (State Standard 1999). The range of this index starts from zero and has no upper border, but this upper border can be set as a maximum possible value (Baranovskiy 2007). To estimate the probability of forest fire danger, the complex meteorological index should be normalised:

$$P_j(D) = \frac{NI_D}{NI_{max}}, \tag{6}$$

where NI_D is a value from the complex meteorological index (Nesterov Index) for the day for which the forecast is realised; NI_{MAX} is the maximum value of the complex meteorological index. The range of variation of forest fire danger from meteorological conditions will then be within the limits from 0 to 1.

The complex meteorological index is calculated by the formula (Kuznetsov and Baranovskiy 2009):

$$NI = \sum_n t(t - r), \tag{7}$$

where t is the air temperature; r is the dew point temperature; n is the number of days after the last rain.

The dew point characterises the amount of moisture in the air. The higher the dew point, the greater the humidity at a given temperature. The dew point temperature is calculated as the temperature to which air must be cooled (at the constant pressure and constant water vapour content) to reach saturation and dew appearance. The saturation state can exist providing that the air contains the maximum possible amount of water vapour at the given temperature and pressure.

Fire danger calculated with the above-stated method differs slightly in different regions within the same forest, and so for clarity, GIS allows visualisation of fire risk maps that allow a quantitative analysis of wildfire risk. It's important to analyse the spatial distribution of wildfire risk in order to delineate and prioritise particularly susceptible areas.

When looking through wildfire risk related literature, one notices a great confusion over the proper use of terminology and the absence of a comprehensive methodology. According to Bachmann (2001) and Bachmann and Allgower (1998, 2001), risk is defined as the product of the probability of fire occurrence (fire danger) and expected negative outcome. The expected negative outcome for a definite forestry region is the total cost of wood located in that region. Information about the number of cubic meters of wood in each region, and its price, is contained in the database. Fire risk assessment is an important part of forest fire prevention which must be made in advance. In contrast to fire danger, which shows what regions are more susceptible to fire occurrence, the risk is a more extensive characteristic and represents not only the most fire-prone areas, but also the areas where the fire will cause the most damage. Fire risk management is a vital element which allows the destructive effects of fires to be eliminated or reduced. For this purpose, our geoinformation system produces fire risk maps through effective spatial data storage and queries.

4 Programme description

The programme realisation of the fire danger assessment system is the web-oriented geoinformation system developed on the Django platform in the programming language Python (Dawson M 2003, Langrangen 2009). The GeoDjango framework was used for realisation of cartographic functions. GeoDjango is a module included in Django that turns it into a geographic web framework. The database management system PostgreSQL, and its spatial extension PostGIS, were used to create a database and to manage the cartographic data. The GeoJSON format was used for encoding collections of simple geographical features along with their non-spatial attributes. The JavaScript library OpenLayers was used for displaying map data in web browsers (Holovaty and Kaplan-Moss 2007; Hourieh 2008, Newman 2008, Tracey 2010). The architecture of the system is depicted in Figure 2.1.

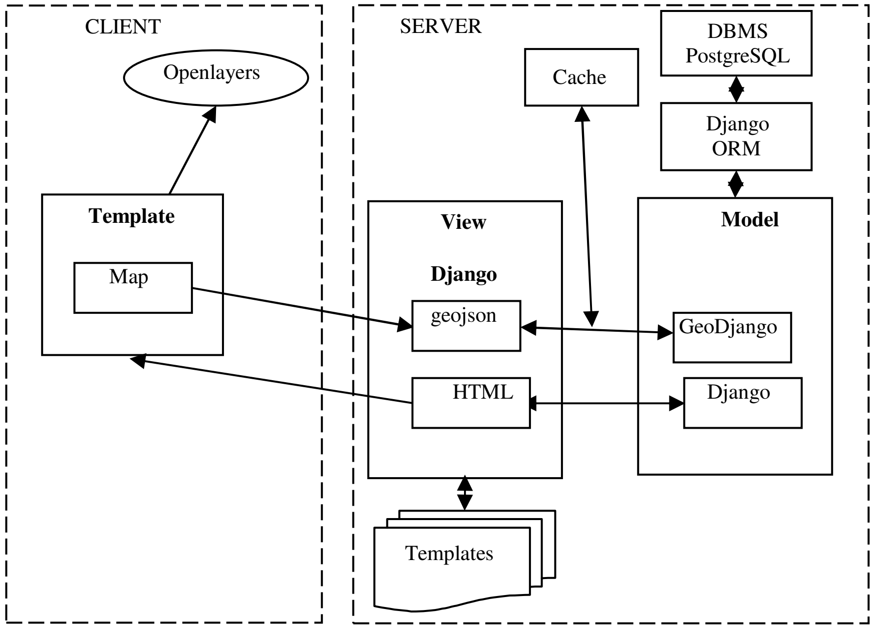


Fig. 2.1 The system architecture

Django's Model View Template (MVT) is quite different from the classic Model View Controller (MVC) architecture where the controller tells the model to change, and the model notifies the view in some way (preferably through a mechanism that allows the view to be fairly decoupled from the model). The Model, View and Controller are all things that exist at the same time in the memory (possibly running in different threads or processes), for extended periods, with their own state, and must interact with each other. Django's Model View Template is quite different from this. In MVT, there is no state, there is only data. For the purposes of most HTTP requests (GET requests), the data in the database is treated as immutable data input, not state (Holovaty and Kaplan-Moss 2007, Hourieh 2008, Newman 2008, Tracey 2010).

It's very important to have visual representation of the most fire-prone areas for forest fire prevention. Maps allow users to highlight places that are susceptible to fires beginning. These maps represent the territory subdivided into different classes depending on the potential risk of forest fire. They reflect an assessment of forest fire risk for each forestry region. Part of a forest fire risk map is depicted in Figure 2.2. Forest fire risk is measured in ten Euros per cubic metre.

Fig. 2.2 Part of a fire risk map

5 Typical results

The software package that runs on the server allows a set of parameters describing forest fire danger to be obtained and visualised. Figure 2.3 demonstrates the graphs. The first graph shows the dependence on temperature in a day of a given month. Values from the Nesterov Index are depicted in the second graph. The final value for forest fire danger is shown in the third graph.

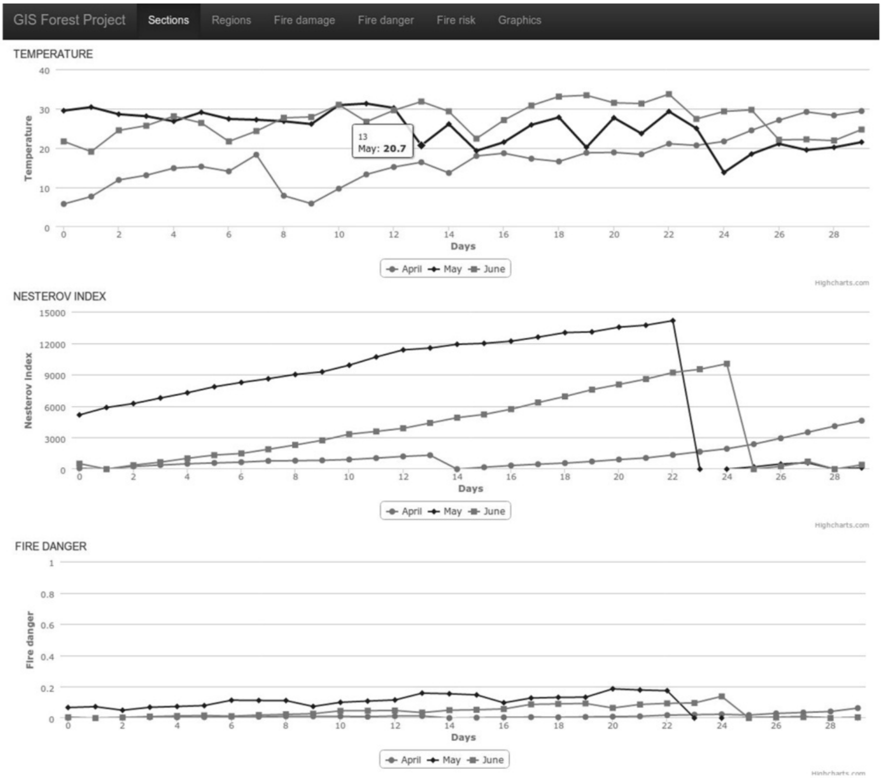


Fig. 2.3 Graphs of dependencies of temperature, Nesterov index and fire danger on a day of three months

It should be noted that remote sensing technologies for forest cover can be used for a more accurate and operative assessment of forest fire danger (Kuznetsov and Baranovskiy 2013). In addition, data from the MODIS Terra/Aqua device (Morisette et al. 2005) can be used to assess the temperature of the atmospheric boundary layer.

6 Conclusion

A new model for fire danger assessment is described in this chapter. This model takes into account meteorological conditions and anthropogenic load. The chapter describes a geoinformation system for forest fire danger prediction based on this model. Knowledge of fire risk areas allows users to recommend the development of fire prevention plans. Fire risk maps, created with the help of geoinformation systems, are considered to be the starting point for a large-scale prevention scheme. The estimation of fire risk helps to identify areas most prone to fire ignition and spread, and to efficiently allocate forest fire fighting resources.

Acknowledgments This work is undertaken with financial support from a state grant with the Ministry of Education and Science within FCP "Researches and developments in priority directions of development of a scientifically-technological complex of Russia on 2007 - 2013". The state contract № 14.515.11.0106.

References

Bachmann AA, Allgower (1998) A framework for wildfire risk analysis. In: Proceedings of the Third International Conference on Forest Fire Research, 14th Conference on Fire and Forest Meteorology; 16-20 November 1998: ADAI Coimbra, Portugal: 2177-2190

Bachmann AA (2001) GIS-based Wildland Fire Risk Analysis. Dissertation, Universidad de Zurich, Zurich

Bachmann AA, Allgower B (2001) Consistent wildland fire risk terminology is needed! Fire Management Today 61(4):28-33

Baranovskiy NV (2007) Technique of forecasting of forest fire danger as a basis of new state standard. Fire Safety 4: 80-84

Dawson M (2003) Python programming for the absolute beginner. Premier Press, USA

Grishin AM (1981) Mathematical modelling of forest fires. Tomsk University Publishing House, Tomsk

Grishin AM, Baranovskij NV (2003) Comparative analysis of simple models of drying of the forest combustibles, including the data of experiments and natural observations. Journal of Engineering Physics and Thermophysics. 76N5: 166-169.

Holovaty A, Kaplan-Moss J (2007) The Definitive Guide to Django: Web Development Done Right. Apress, New York

Hourieh A (2008) Learning Website Development with Django. Packt publishing, Birmingham

Khodakov VE, Zharikova MV (2012) Forest Fires: Research Methods. Grin DS, Kherson

Kuznetsov GV, Baranovskiy NV (2009) Forecast of Forest Fire Occurrence and their Ecological Consequences. Publishing house of the Siberian Branch of the Russian Academy of Science, Novosibirsk

Kuznetsov GV, Baranovskiy NV (2013) Focused sun's rays and forest fire danger: new concept. Proceedings of SPIE – The International Society for Optical Engineering. Vol. 8890, paper 889011; doi:10.1117/12.2033929

Langrangen HP (2009) A Primer of Scientific Programming with Python. Springer, Oslo

Morisette JT, Giglio LI, Justice CO (2005) Validation of the MODIS active fire product over Southern Africa with ASTER data. International Journal of Remote Sensing. Vol. 26. N 19:4239-4264

Nazarov AA, Terpugov AF (2006) Probability Theory and Stochastic Processes: Manual. Publishing House NTL, Tomsk

Newman S (2008) Django 1.0 Template Development. Packt publishing, Birgminham

Simard AJ (1976) Wildland Fire Management: the Economics of Policy Alternatives. Forestry Technical Report 15, Department of Fisheries and Environment, Canadian Forestry Service, p 52

State Standard 22.1.09-99 (1999) Monitoring and forecasting of forest fires. General requirements. Moscow: State Standard of Russia

Tracey KM (2010) Django 1.1. Testing and Debugging. Packt publishing, Birgminham

An Integrated Web Mapping Solution to Assess the Effect of SLR on the Northern Coast of EGYPT – EGSLR

Aymen A. Solyman, Mamdouh M. El-Hattab, Rifai I. Rifai and Waleed M. T. Fakhar[1]

Abstract In its 2007 assessment, the Intergovernmental Panel on Climate Change (IPCC) declared the Nile Delta one of three sites on Earth that are most vulnerable to sea level rise (SLR). The Panel projected that the global average surface temperature will increase by 1 to 3.5C with an associated rise in sea level of 15 to 95cm by 2100. Several recent assessments suggest this figure could be much higher. In one study that considered the impact of a 1m SLR for 84 developing countries, Egypt was ranked the 2nd highest with respect to the coastal population affected, 3rd highest for coastal GDP affected and 5th highest for proportion of urban areas affected. In this chapter we present a methodology to build an integrated web map solution (EGSLR) to access the effect of the SLR scenarios on the northern coast of Egypt. EGSLR covers the area from Alexandria to Port Said which extends to about 285 kilometres along the Mediterranean coast. The overall objective of the EGSLR is to develop a web vulnerability mapping solution to study the different sea level rise scenarios and to identify and map the areas that are most vulnerable to SLR on the Nile Delta coast. Four scenarios are implemented to assess the impact of SLR, for 25 cm, 50 cm, 75 cm and 100 cm.

1 Introduction

The northern, or the Mediterranean coast of Egypt extends from Alexandria city in the west to Port Said city in the east, with a total length of about 240 kilometres. The upper left (UL) corner of the study area lies at latitude: 31.8, longitude: 29.6,

---

A. A. Solyman, M. M. El-Hattab, R. I. Rifai
Environmental Studies and Research Institute,
University Of Sadat City
asolyman@cedare.int

Waleed M. T. Fakhar
Faculty of Information and Computing, Arab Academy
Cairo

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_3,

and the lower right (LR) corner is at latitude: 30.8, longitude: 32.6.The Mediterranean coastal zone is the major industrial, agricultural, and economic resource of the country. The Nile delta and Mediterranean coast comprise 30-40% of Egypt's agricultural production, and half Egypt's industrial production. The three main delta lagoons, Idku, Burullus and Manzala, produce over 60% of Egypt's fish catch. Approximately 15% of Egypt's GDP is generated in these Low Elevation Coastal Zone (LECZ) areas. This area is at risk from future sea level rise and storm surge. With a large and growing population in coastal zones and a low adaptive capacity, the area is highly vulnerable. In the absence of adaptation, the physical, human and financial impacts of climate change on coastal zones will be significant. Coastal adaptation is therefore likely to be a priority.

2 Global Projected Sea Level Rise

Projections of SLR have changed over the years as more information has become available (e.g., more advanced climate change models and more accurate data).The 1990 IPCC reported a scenario of global warming and consequent global SLR of 18 cm by 2030 and between 21 cm to 71 cm by 2070 (IPCC 1990). In 2001, the IPCC projected that SLR would increase by 9cm to 88cm over 1990 sea levels by 2100 (IPCC TAR 2001). The range of SLR predictions is due to uncertainties about greenhouse gas emission scenarios, the temperature sensitivity of the climate system, contributions from the Antarctic, and glacial melt.

Table 1 Projected temperature change and sea level rise (excluding rapid dynamic future changes in ice flow) for the six IPCC emission scenarios

IPCC emission scenario	Temperature rise IPCC-AR4	Sea level rise 2100 IPCC-AR4	Sea level rise 2050 IPCC-TAR
B1	1.8C	0.18 - 0.38m	0.05 – 0.26m
A1T	2.4C	0.20 - 0.45m	0.07 – 0.29m
B2	2.4C	0.20 – 0.43m	0.06 – 0.28m
A1B	2.8C	0.21- 0.48m	0.06 – 0.28m
A2	3.4C	0.23 – 0.51m	0.06 – 0.27m
A1FI	4.0C	0.26 – 0.59m	0.06 – 0.3m

In February 2007(AR4), the IPCC slightly lowered its estimate of SLR to between 18 cm to 59 cm over 1990 sea levels by 2100, because new data and technologies became available about the contribution of the thermal expansion of SLR, but the new range does not incorporate the potential acceleration of the melting of

Greenland or the West Antarctic Ice Sheet. The estimates include only the steric component of the sea level rise due to the heating of the ocean waters and their consequent expansion. The numbers given by IPCC should therefore be considered as a lower limit of expected sea level rise. The 2007 IPCC report advises that if the ice discharge from these processes were to increase in proportion to global average surface temperature change, it would add 0.1-0.2 m to the upper limit of sea level rise by 2100. The IPCC also acknowledged that “larger values could not be excluded” as shown in Table 1.

3 The Methodology of the Development of Interactive Web Enabled SLR Scenarios (EGSLR)

The most important factors being included in the design and implementation of EGSLR as a GIS application for the internet are functional modules, a graphic user interface, download time and system performance. The system must be portable and extendable to accommodate future changes in hardware, software and networking. With this is mind, EGSLR was developed based on two components: the server side and the web client side, which run in the browser. The framework of the technical development process of the EGSLR focuses on:

1. Spatial and attribute database collection, analysis and design
2. Data conversion and import
3. Development environment and tools
4. EGSLR database model
5. EGSLR architecture
6. EGSLR main features
7. EGSLR main components
8. EGSLR main functions

4 EGSLR Database Model

The database is the most important part of any web application. In the case of GIS web-based applications, it becomes more important because of the storage requirements of the spatial data. Once the spatial data is stored in a database it can be used, analysed and displayed in the form of maps by a web-based application. EGSLR data is divided logically into two categories: spatial data and attribute (tabular) data as shown in Figure 3.1. The EGSLR database model is designed to be stored physically in different format although the relationships between the two categories of data are preserved regardless of whether the division is physical or logical (in a hybrid database model). This means that two separate databases were used in the system; one for spatial data and one for attribute data.

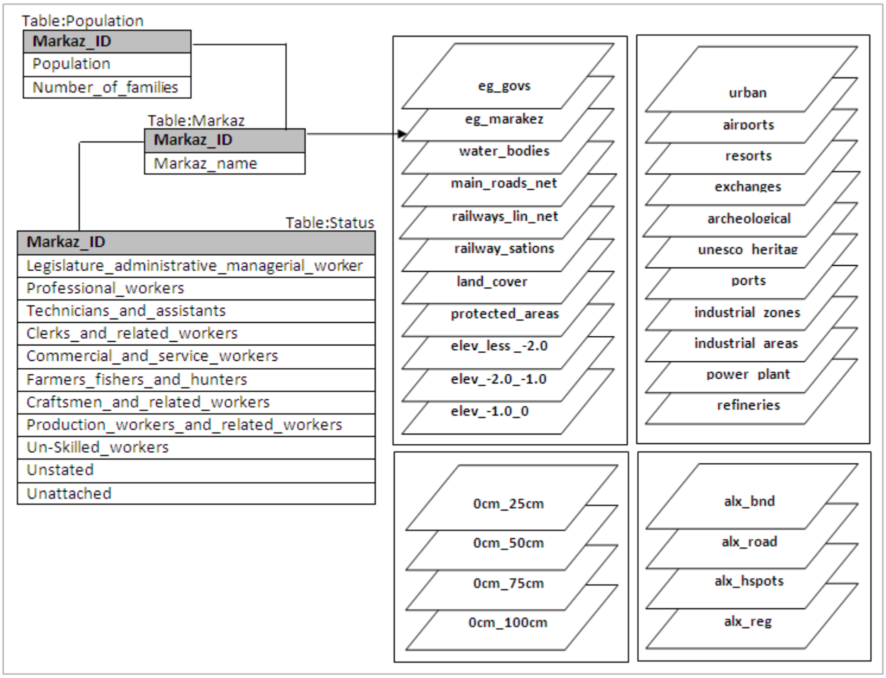


Fig. 3.1 EGSLR database model

5 EGSLR Architecture

There are basically two types of architecture for developing web-based GIS applications: client-side, and server-side. In a client-side GIS web application, the client (web browser) is enhanced to support GIS functionality, while in a server-side GIS application, a web browser is used to generate server requests and display the results. EGSLR is a server side application (with a thin client architecture) where users can send a request to a server (i.e. an address), and the server processes the request and sends the results back as an image embedded in an HTML page via standard HTTP (HyperText Transfer Protocol). The response is shown as a standard web page that a generic browser can view in server-side web GIS applications; all the complex and proprietary software, in addition to the spatial and tabular data remain on the server. This architecture has several advantages, because the application and data are centralised on a server.

6 EGSLR Main Components

EGSLR is divided into a set of functional units (modules). Each module represents a set of related tasks or functions. The modules are independent of one another but they do communicate with each other. EGSLR is developed and deployed using standard web development tools, and is comprised of two elements, the web site framework and the functional tools. The framework presents the EGSLR supporting information to the user via a graphical user interface. The second element are the functional tools that enable access to GIS functions such as SLR scenarios, mapping and query functions. The main modules are:

1. Display functions - map tools
2. Cartographic presentation
3. Utilities
4. Mashup module
5. Layer manager module
6. Sea level rise scenarios module: This module contains four levels to study:
   - Scenario 1: sea level rise from 0 cm to 25 cm.
   - Scenario 2: sea level rise from 0 cm to 50 cm.
   - Scenario 3: sea level rise from 0 cm to 75 cm.
   - Scenario 4: sea level rise from 0 cm to 100 cm.
7. Population distribution module

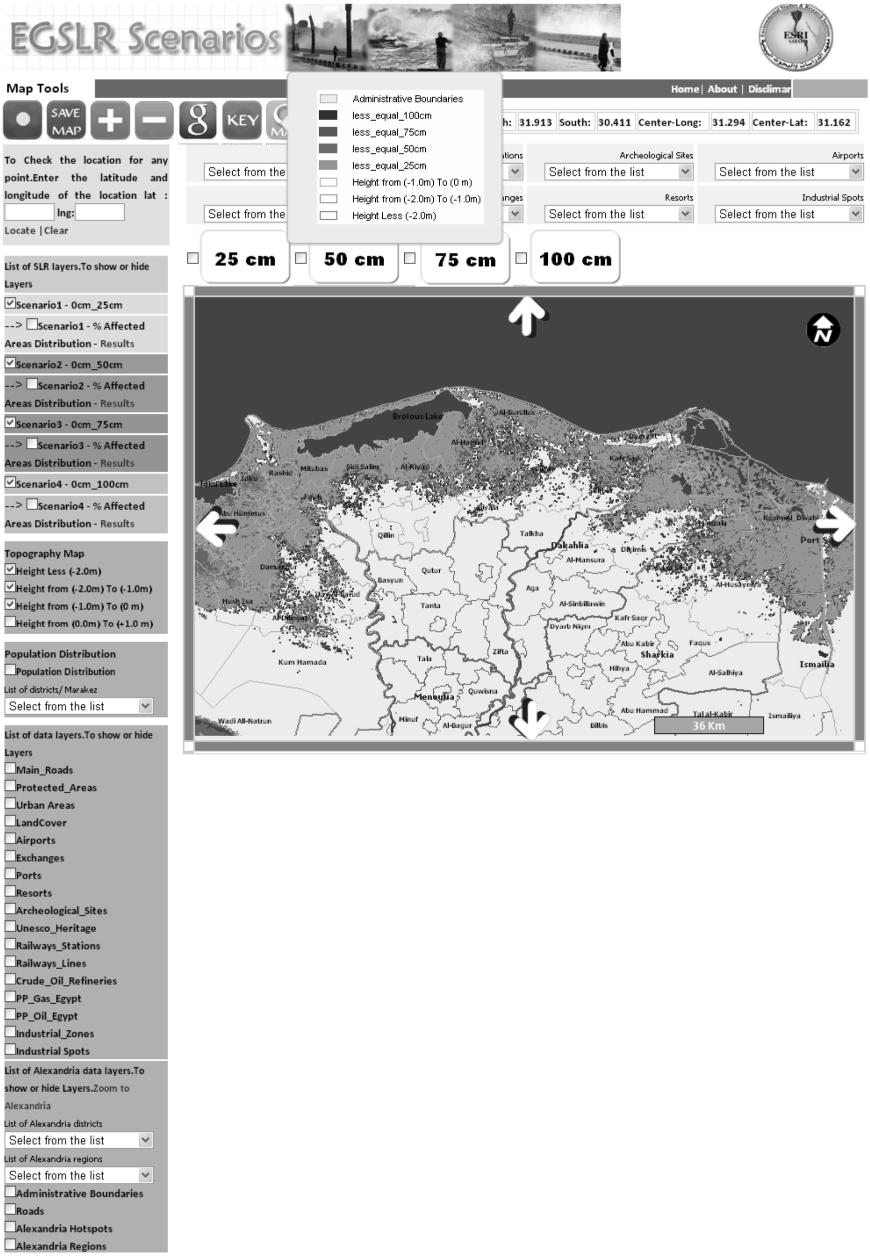


Fig. 3.2 The main interface of the EGSLR (www.egslr.com)

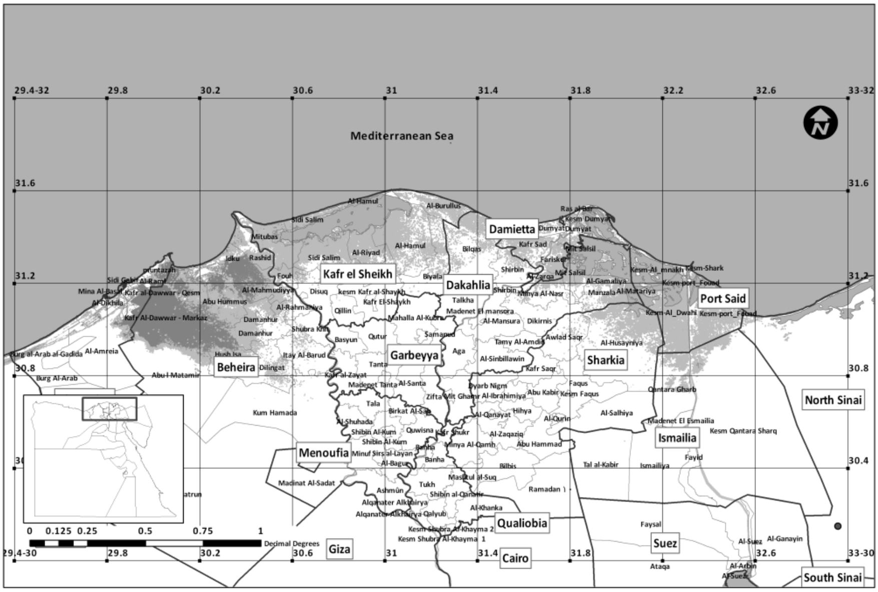


Fig. 3.3 Scenario (2) the expected impact of a 50cm SLR

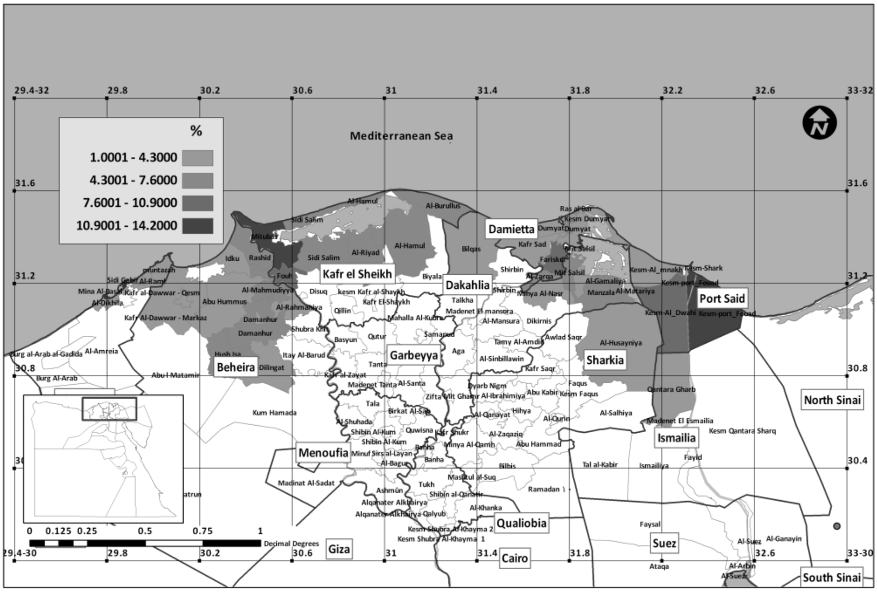


Fig. 3.4 Map shows the classification of impacted districts according to Scenario 1 – 25 cm

7 Conclusions

1. The Nile Delta shoreline contains high population densities, significant economic activities and ecosystem services. This area is already subject to climate change, which has the potential to pose increasing risks to this coastal zone in the future. However, the effects of climate change need to be seen in the context of other socioeconomic drivers.
2. Global mean atmospheric temperature rise is assumed to be the main driving force for sea level rise. The most recent IPCC assessment based on the most gloomy scenario puts predictions of 21st century sea level rise at between 26 and 59cm (10-23 inches).
3. The high risk areas include parts of Alexandria, Behira, Dakahlia, Damietta, Ismalia, Kafr Alshikh, Port Said and Sharkia governorates.
4. The most vulnerable governorates are Kafr Alshikh, Behira and Port Said governorates which present about 70 to 75% of the total affected area in each scenario.

References

Dasgupta S, Laplante B, Meinsner C, Wheeler D and Yan J (2009) The impacts of sea level rise on developing countries: A comparative analysis. *Climate Change*, vol 93, pp 379-388

El-Raey M (1997) Vulnerability assessment of the coastal zone of the Nile delta, Egypt, to the impacts of sea level rise. Ocean and coastal management, vol 3, pp 29-40

El-Raey M, Dewidar Kh and El-Hattab M M (1999) Adaptations to the impacts of sea level rise in Egypt. journal of mitigation and adaptation strategies for global change, vol 4, pp 343-361

Elsharkawy H, Rashed H, and Rached I (2009) Climate change: the impacts of sea level rise on Egypt, 45th international society of city and regional planners (ISOCARP) Congress

Frihy O (2003) The Nile Delta-Alexandria coast: vulnerability to sea-level rise, consequences and adaptation. mitigation and adaptation strategies for global change, vol 8, pp 115-138

Gangfeng M A and Shuguang LIU (2003) A forecasting model of delta evolution using artificial neural networks. International conference on estuaries and coasts, China, vol 1, pp 841-846

Gebert N, Kloos J and Birkmann J (2012) Emerging Risks: sea level rise and potentially forced and planned relocation - Case study from Greater Alexandria, Egypt, Institute for Environment and Human Security, United Nations University, Final Report of CLICO project

IPCC Climate Change (1990) The IPCC scientific assessment report, Cambridge university press, Cambridge, UK

IPCC Climate Change (2001) The scientific basis, Contribution of Working Group 1 to the third assessment report of the Intergovernmental Panel on Climate Change (IPCC)

IPCC Climate Change (2007a) The physical science basis, working group 1 contribution to the fourth assessment report of the Intergovernmental Panel on Climate Change (IPCC), technical summary.

IPCC Climate Change (2007b) Mitigation, working group III contribution to the fourth assessment report of the Intergovernmental Panel on Climate Change (IPCC), technical summary

Rahmstorf S, Foster G and Cazenave A (2012) Comparing climate projections to observations up to 2011, Environmental Research Letters, vol 7, pp 1-6

Rabus B, Eineder M and Bamler R (2003) The shuttle radar topography mission- A new class of digital elevation Models acquired by space borne radar. Journal of photogrammetry and remote sensing, vol.57, pp. 241-262

Vermeer M and Rahmstorf S (2009) Global sea level linked to global temperature. Proceedings of the National Academy of Science of the USA, vol 106, pp 21527-21532

ISOCARP Congress. http://www.isocarp.net/Data/case_studies/1456.pdf. Accessed 1 June 2013

Heattile, a New Method for Heatmap Implementations for Mobile Web-based Cartographic Applications

Sebastian Meier, Frank Heidmann, Andreas Thom[1]

Abstract Mobile handheld computing devices are becoming a more important part of our digital infrastructure. Web-based map services and visualisations for spatial data are being optimised for use on these new mobile devices. Despite the fast development of mobile devices, compared to laptop and desktop computers they remain slower in terms of processing power and have limited memory capacity. Bandwidth in most areas is still limited, regardless of the fast expansion of high-speed internet for mobile devices. To overcome these two limitations an improvement in existing spatial data visualisations is necessary. This chapter compares existing methods and presents a new approach to generating and delivering web-based heat map visualisations for spatial data that is optimised for mobile devices.

1 Introduction

Heat maps, also known as density maps, density surfaces and shaded isarithmic maps, are usually used for visualising evenly distributed spatial data points with varying values (e.g. weather maps with varying temperatures) or spatial data-point clusters representing density (e.g. network-coverage-maps with one data point per antenna). With the help of heat maps a user can identify hotspots, distribution, and correlations within a spatial dataset (Pettit et al. 2012). Through further development of web technologies these visualisations have also become of interest for web-based geo-visualisations.

Since the introduction of the iPhone and the beginning of the still growing smartphone segment, maps, or rather navigation, has been an essential part of the applications (apps) available. Today, there are numerous apps, map providers, navigation providers and many commercial and open source tools, which allow

S. Meier, F. Heidmann, A. Thom
Interaction Design Laboratories
University for Applied Science
Potsdam, Germany
e-mail: meier@fh-potsdam.de

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_4,

users and developers to bring maps to mobile devices. In addition to maps and navigation, geovisualisations (visualisations of spatial data), are also starting to appear on mobile devices. Although they have been present on the internet for years, the existing approaches to geovisualisations have to overcome a number of technological and conceptual obstacles until they can fully embrace the new mobile medium. This chapter concentrates on the *heat map* geovisualisation, and a new method to overcome the obstacles described above.

2 Existing Methods

The existing approaches to generating and delivering heat maps for web-based map systems can be categorised into two methods. The first method, referred to here as the *client-side method,* delivers raw spatial data, depending on the technology used, in JSON, XML, CSV or other similar formats, to the user's browsers. The browser then transforms this data, using client-side technology such as JavaScript, into the actual heat maps (D3 2014, Heatmapjs 2014, Heatcanvas 2014, WebGL 2014, Google1 2014).

The second method, referred to here as the *tile-based method*, involves prerendering the heat maps and then storing the visualisation data in tiled images on a server. The image-data is then delivered to the user's browser in a standard tile-based system used by most common web map systems (CartoDB 2014, Nokia 2007).

3 Problem Definition

The existing approaches work very well in stationary browser applications and currently, in 2013, there are already various commercial and open source implementations of these methods. When we try to use the existing methods on mobile handheld web browsers, however, there are two major difficulties. One is the *limited performance factor* of mobile handheld devices. Even though the latest developments have turned what were not long ago still mostly text-input devices into small powerful computers, they are still not comparable to laptops or desktop computers. This limiting factor is a problem more relevant for the *client-side method*. Storing a large amount of data for real time use in the browser can be difficult on mobile devices and is limited, depending on the device. Rather seriously, this involves generating heat maps in real time, because this process is computationally intensive and also limited.

The second major limitation is the *bandwidth*. Even with the implementation of the latest network technology the bandwidth in rural areas and densely populated areas is still limited. As a result, the *tile-based method* is problematic because the delivery of heat maps through the image data used by the *tile-based method* is very bandwidth intensive.

In addition to these two major limiting factors there are two additional factors involving the technological features of modern handheld devices. On the one hand, we have more and more *high-resolution displays* being built into modern devices. These high-resolution displays require the *tile-based method* to deliver even bigger images, making it even more bandwidth-intensive. On the other hand, multitouch displays and modern web-map systems allow a user to smoothly zoom in and out as well as interact with the data visualised on screen. While zooming is no problem for the *tile-based method,* further interaction is limited. In the *client-side method* interaction is possible, as is zooming, even though zooming can be very resource consuming.

Table 4.1: Comparison of existing heat map methods

	Tile-based Method	Client-Side Method
Computationally intensive	Low	High
Bandwidth intensive	High	Medium
Interactivity	Low	High
High-Resolution	Yes, but requires more bandwidth intensive images	Yes, Vector-Data
Real time	No, due to pre-rendering	Yes

As shown above, according to the comparison of existing approaches on mobile web browsers, a new method should have: 1. A *small data footprint* regarding bandwidth, 2. require *as little client-side calculation as possible*, as well as being 3. a *high-resolution visualisation* and it should also enable, *4. interaction with the visualised data* itself.

4 The New Heattile Approach

The new approach consists of three steps. The first step is a raster-based clustering and generalising of the data, which is then rendered on a request-basis into GeoJSONs (GeoJSON 2014), delivered to the browser and then visualised client-side for further interaction.

4.1 Server-Side Pre-Clustering

For the following steps we assume that the data is stored in a database. The sample queries are MySQL-Queries. The first step is clustering the spatial data. We use the tiling approach, which was initially introduced through the WMS Specification (Opengis1 2014) by the Open Geospatial Consortium (OGC 2014) in 2000. It became popular through modern map providers and their implementation of the Web Map Tile Service Implementation Standard (Opengis2 2014), such as that of Google (Google2 2014), Bing (Bing 2014), OpenStreetMap (OpenStreetMap 2014) and Mapbox (Mapbox1 2014). The tiling approach is using a web-Mercator projection, which results in an orthogonal and evenly distributed coordinate system (see Fig. 4.1 right). This coordinate system is then split into smaller squares, using a Quad-Tree-Method (Potmesil 1997), starting at one square at zoom-level x, four squares at zoom-level x+1, sixteen squares at zoom-level x+2 resulting in zoom-level x, tile number t = 2^((x-1)*2). In each zoom-level every square has a unique identifier. Starting at zero in the upper left corner, row after row, every column receives the consecutive number as an identifier (see Fig. 4.2).

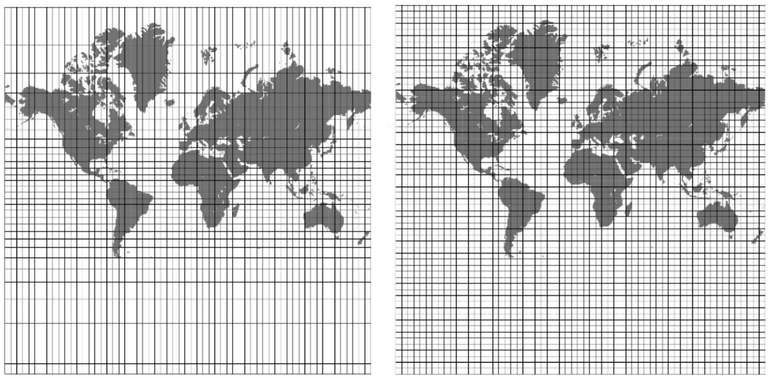


Fig. 4.1 Coordinate-system comparison

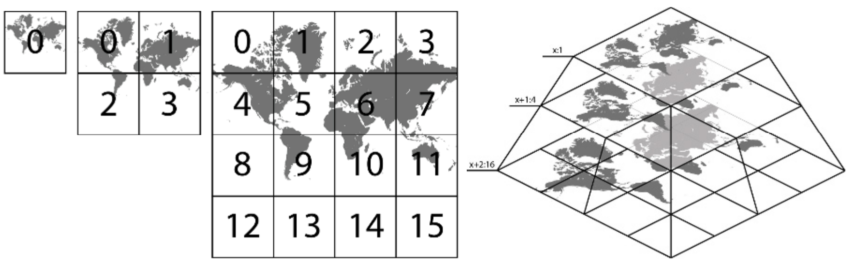


Fig. 4.2 Tiling-System for Spatial Data

This process will be applied to our spatial data; every data point receives an id for every zoom-level, allowing us to receive all data points for a specific tile-id at a specific zoom-level through a very simple query. This technique is inspired by the hierarchical clustering process (Delort 2010).

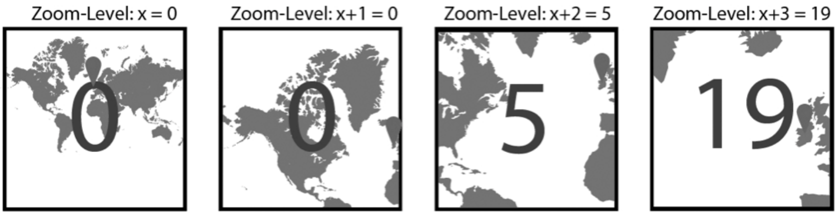


Fig. 4.3 Grouping data points on a per-tile basis: Tile-ID per zoom-level

1. [SELECT GROUP(*), * FROM geo_data WHERE z_(ZOOM_ID) = (TILE_ID) GROUP BY z_(ZOOM_ID)]

After the initial conversion process, the same process of generating tile ids for spatial data points can be applied to new points whenever they are added without touching the already clustered data.

4.2 Server-Side GeoJSON

In order to keep a small data footprint, this method pre-clusters the data into the tiles generated in the step above. Instead of delivering all data points, the method only delivers a number of data points per tile. Instead of returning all tiles for the current zoom level, the method only returns the tiles for the currently visible area on the mobile device plus an additional margin around the requested area. The extra margin gives the user smoother interaction when panning the “slippy map” (Slippy Map 2014). Using this approach, developers can decide how much data they want to deliver to the user’s device, either by shrinking or expanding the margin, and by defining the size of the tiles being displayed at each zoom-level.

To keep the amount of rendering as small as possible, the method does not deliver raw data that needs to be turned into a visualisation on the client’s side, but instead delivers GeoJSONs. The method delivers a range of GeoJSONs, which can be defined by the developer. In the example visualisations in Figures 4.4 and 4.5 we used an evenly distributed range from 0 to 9, 0 holding the tiles with the least data points and 9 the tiles with the most data points. For faster range-computation we recommend caching the maximum value of data points per tile for every zoom-level.

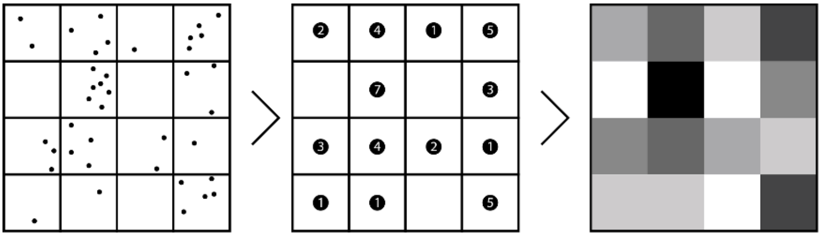


Fig. 4.4 Server-side rendered GeoJSON-Layers

2. [SELECT GROUP(*), * FROM geo_data WHERE z_(ZOOM_ID) = (TILE_ID) WHERE latitude < (LATITUDE_MAX+MARGIN) AND latitude > (LATITUDE_MIN-MARGIN) AND longitude < (LONGITUDE_MAX+MARGIN) AND longitude > (LONGITUDE_MIN-MARGIN) GROUP BY z_(ZOOM_ID)]

4.3 Client-Side Rendering & Interaction

We used Leaflet (2014), a common web-mapping framework, in the user's browser for displaying a map and requesting the GeoJSON-layer with the heat map for the currently visible area. On the client-side we can decide on visual features of the GeoJSON, manipulate those features, and add interactions for each range step (e.g. click or double-click). The colours can, for instance, be modified on the client: in the example below we used higher opacity for layers with higher density. The mono-colour theme was used instead of the common blue to red colour range for heat maps, as the mono-colour range performs better in geovisualisations (Borland 2007, Harrower and Brewer 2003).

Fig. 4.5 Two different heat map visualisations on top of maps

5 Advantages and Disadvantages

The data footprint for the tiled GeoJSONs is very small (see Table 4.2), the processing-power required is little (see Table 4.2) and all processes can be done in real time. The GeoJSONs are visualised as vectors, which gives us high-resolution visualisations. Interaction in our method is limited; we can differentiate between interactions on the layers below the heat map and interactions on each range step, but what we cannot do is receive interactions on a specific tile, which would require splitting the GeoJSON into one GeoJSON per tile, which would result in a higher data footprint as well as more processing. As a work-around we can translate the latitude and longitude values of the interaction into a tile-id and request further information from the server. Another disadvantage is the limitation in terms of visualisation styles. Due to the clustering method used, we are limited to square-based visualisations. For additional visualisations we have created a slightly altered method: every even row is offset by half the square size to the right, allowing us to create other types of visualisations, (Fig. 4.6).

Fig. 4.6 Alternative offset grid with hexagon heat map visualisation

6 Comparison

It must be noted that the amount of data used influences the performance of the three methods. In this chapter we used datasets including from 200 to 200,000 data points for the central area of Berlin, for the testing described below (Table 4.2-4.4), and we used a subset of 14,864 data points. The testing device was an iPhone 5s.

In a detailed comparison (Table 4.2) we can see that the GeoJSONs from our new method are smaller than in the other two data formats. For the image tiles, there is no information available for the uncompressed size as the tiles are compressed on the server. Execution time does not include loading time. The time for the raw data approach can be split into 20-30ms for parsing the data into the visualisation system and an additional 480-570ms for the rendering process. Parsing time is time required for JavaScript to go through the received JSON file and pass it on to the visualisation functions. The time for the raw data was calculated using the Heatcanvas (2014) library. The other libraries mentioned in Section 2 (D3 2014, Heatmapjs 2014, WebGLHeatmap 2014, Google1, 2014) were also tested, and we received varying results (all >200ms) for the rendering time; the parsing time in all the libraries used was nearly the same (20ms-40ms).

While the tile-based method is only limited by the capacity of the pre-rendering process and the new method is only limited by the performance of the database that holds the data, the performance of the client-side method, as described in Section 3, is extremely dependent on the size of the dataset. While performing well on very small datasets, the bigger the datasets become, the worse the performance (see Table 4.4), which at some point makes the client-side method unusable.

Table 4.2 Filesize and execution time of GeoJSONs, varying in tile size, compared to image tiles and raw data for a 500x500px region

Size in Kilobytes	GeoJSON (x)	GeoJSON (x+1)	GeoJSON (x+2)	Image Tiles (lowres)	Image Tiles (highres)	Raw Data
Size	14,4	35	92,1	n/a	n/a	544
Size GZIP-compressed	1,3	5,1	13,5	159,1	430,7	51,8
Execution time in ms	7-9	8-10	9-11	10	15	20-30 (500-600)

Table 4.3 Comparison of the existing methods with the new method

	Heat Tile Method	Tile-based-Method	Client-Side-Method
Computationally intensive	Low	Low	High
Compare to Table 4.2			
Bandwidth intensive	Low	High	Medium
Compare to Table 4.2			
Interactivity	Medium	Low	High
High-resolution	Yes, vector data	Yes, but requires more bandwidth intensive images	Yes, vector data
Real time	Yes	No, due to pre-rendering	Yes

Table 4.4 Execution Time for the client-side method with raw data for a 500x500px region

Data points	14864	29728	44592
Execution time in ms	20-30 (500-600)	44-50 (950-1050)	65-80 (1430-1550)

If we compare the new method with the existing methods, the new method performs better or equally as well in every category, except interactivity. The only disadvantage that remains is the variety in terms of visualisation and interaction.

7 Application and Future Works

There is a wide field of applications for the new method. In the MSNI research project (MSNI 2014), we used the method for displaying the density of different types of locations within a city (e.g. restaurants, bars, museums, etc.) which allows the user to explore a city while being on the road (similar attempts have been made by Nokia (2007)). In a similar use case researchers developed a visualisation for social-network data from foursquare to inform tourists about nearby locations and activities (Komminos et al. 2013). Another possible use case is the visualisation of environmental data on mobile devices to allow users to explore their environment while on the move.

To overcome the persistent obstacles we tried using the clustered data as raw data for existing JavaScript-based heat map libraries (Fig. 4.7). This resulted in even smaller data footprints, as we were able to strip the GeoJSON data and turn it into simple JSON (JSON2 2014) files. On the other hand it also resulted in more client-side processing, which slowed the visualisation.

In the future we plan to apply new methods of shrinking the file-size of the GeoJSONs even further through compression (Hanov 2010, json.hpack 2014, JSON1 2014) or the BSON (2014) method. Additionally we need to look into caching systems for further optimisation and better scalability of our approach, similar to the techniques already used in the tile-based method (Liu et al. 2007).

The latest work on vector-tiles, for example that by Google Maps (Google2 2014) or MapBox (MapBox2 2014), have improved the data footprint of the tile-based-approach. If the new vector method by Google and MapBox can be applied to the tile-based method for heat maps this could improve the limitation in terms of bandwidth, although it will not improve the ways a user can interact with the visualisations generated by this method.

Fig. 4.7 Using the method to deliver clustered raw data for client-side heat map libraries

8 Conclusion

The method described in this chapter presents a useful alternative to existing approaches for visualising heat maps on mobile handheld devices. The new approach performs better in the major categories we identified as obstacles for geovisualisations on mobile devices. The fact that the generation of heat maps can be done in real time is a particularly interesting new possibility for applications. The approach also performs well in the other categories. The only downside is the variety of visual representations that it is possible to achieve, due to the raster-based clustering.

The code required for using the method is available on GitHub under GPL/MIT: https://github.com/sebastian-meier/HeatTiles.

References

Bing (2014) http://maps.bing.com, Accessed: 10.01.2014

Borland D, & Taylor MR (2007) Rainbow color map (still) considered harmful. IEEE Comput Graph Applications, 27(2), 14-17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17388198, Accessed: 10.01.2014

BSON (2014) BSON specification. http://bsonspec.org Accessed: 10.01.2014

CartoDB. http://cartodb.com/visualise Accessed: 10.01.2014

D3 (2014) D3 – Hexbinding. http://bl.ocks.org/mbostock/4330486 & http://bl.ocks.org/mbostock/4248145 Accessed: 10.01.2014

Delort JY (2010) Hierarchical cluster visualization in web mapping systems. In Proceedings of the 19th international conference on World wide web (WWW '10). ACM, New York, NY, USA, 1241-1244. DOI=10.1145/1772690.1772892

GeoJSON (2014) http://www.geojson.org Accessed: 10.01.2014

Google1 Heatmaps. https://developers.google.com/maps/documentation/javascript/examples/layer-heatmap, Accessed: 10.01.2014

Google2 (2014) Maps. http://maps.google.com, Accessed: 10.01.2014

Hanov S (2010) Compress JSON with automatic type extraction. URL http://stevehanov.ca/ blog/index.php?id=104 Accessed: 10.01.2014

Harrower M, Brewer CA (2003) ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. Cartographic Journal, 40(1), 27-37. doi:10.1179/000870403235002042

Heatcanvas (2014) https://github.com/sunng87/heatcanvas Accessed: 10.01.2014

Heatmapjs (2014) http://www.patrick-wied.at/static/heatmapjs/ Accessed: 10.01.2014

JSON1 (2014) Compression Algorithms. http://web-resource-optimisation.blogspot.com/2011/06/json-compression-algorithms.html Accessed: 10.01.2014

JSON2 (2014) http://json.org/ Accessed: 10.01.2014

json.hpack (2014) https://github.com/WebReflection/json.hpack/wiki Accessed: 10.01.2014

Komninos A, Besharat J, Ferreira D, Garofalakis J (2013) HotCity: enhancing ubiquitous maps with social context heatmaps. In Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM '13). ACM, New York, NY, USA, , Article 52 , 10 pages. DOI=10.1145/2541831.2543694

Leaflet (2014) http://www.leafletjs.com Accessed: 10.01.2014

Liu Z, Pierce ME, Fox GC, Devadasan N. (2007) Implementing a caching and tiling map server: a Web 2.0 case study (pp. 247–256). Presented at the 2007 International Symposium on Collaborative Technologies and Systems (CTS), IEEE. doi:10.1109/CTS.2007.4621762

Mapbox (2014) http://www.mapbox.com Accessed: 10.01.2014

Mapbox (2013) Vector Tiles for Mapbox Streets. https://www.mapbox.com/blog/vector-tiles, Accessed: 10.01.2014

MSNI (2014) http://www.m-s-n-i.de, Accessed: 10.01.2014

Nokia (2007) Identifying interesting locations based on commonalities in location based postings. http://www.google.com/patents?vid=USPATAPP11784086, Accessed: 10.01.2014

Open Geospatial Consortium (2014) http://www.ogc.org Accessed: 10.01.2014

Opengis1 (2014) Web map service (WMS) implementation specification. http://www. opengeospatial.org/standards/wms Accessed: 10.01.2014

Opengis2 (2014) Web Map Tile Service Implementation Standard: http://www.opengeospatial.org/standards/wmts Accessed: 10.01.2014

OpenStreetMap (2014) http://www.openstreetmap.org Accessed: 10.01.2014

Pettit C, Widjaja I, Russo P, Sinnott R, Stimson R, Tomko M (2012) Visualisation support for exploring urban space and place. In XXII Congress of the International Society for Photogrammetry and Remote Sensing, Melbourne, Australia

Potmesil M, Bell Laboratories, Lucent Technologies (1997) Maps Alive: Viewing Geospatial Information on the WWW. Sixth International World Wide Web Conference. April 7-11, 1997 Santa Clara, California USA.

Slippy Map (2014) http://wiki.openstreetmap.org/wiki/Slippy_Map Accessed: 10.01.2014

WebGLHeatmap (2014) http://ursudio.com/webgl-heatmap-leaflet/ Accessed: 10.01.2014

Part II
User-Oriented Map Design and Production

Designing Simple and Complex Animated Maps for Users from Different Age Groups, Employing the Appropriate Selection of Static and Dynamic Visual and Sound Variables

Dariusz Dukaczewski[1]

Abstract The appropriate selection of static and dynamic visual and sound variables (and related methods of cartographic presentation) is one of the vital factors for proper understanding of geographic processes, especially when dealing with animations that are intended to be read by users of differing age groups. In the research reported here, the author analysed and synthesised the available information related to user age, generation and education level that might have affected the visual and aural perception of variables. This was used as the basis to propose guidelines concerning the use of static and dynamic visual and sound variables, as well as methods of cartographic presentation for eight age groups. The utility of these guidelines was then tested on groups of representatives, enabling the rules to be further refined. Based on these modified guidelines, tables of rules were proposed for the use of variables and methods for each age group, which were integrated into the author's *entities–cartotrophic method* of designing simple animations and *entities–polystaymic method* of designing complex animations.

1 Introduction

There is no doubt that after over 50 years of development of simple animated maps, and 30 years of development of complex animated maps, both types have become (from a technical point of view) operational tools for the visualisation of dynamics. However, notwithstanding the achievements of many years of work, the process by which these maps are designed still encounters a considerable number of obstacles, most of these being related to the need to comply with methodological conditions, as well as perceptions of these animations.

Research into the cartographic transfer of information through cartographic animations carried out by Peterson (1998), Schnotz et al (1998), Midtbø (2003),

D.R. Dukaczewski
IGiK - Institute of Geodesy and Cartography, Warsaw, Poland
e-mail: dariusz.dukaczewski@igik.edu.pl

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_5,

Opach (2005), Hanewinkel and Tzchaschel (2005), as well as the quantitative research on the efficiency of this process by Koussoulakou and Kraak (1992), Harrower et al (1998, 2000), Lowe (1999, 2003, 2004), Bétrancourt et al (2000), Slocum et al. (2001), Fabrikant and Goldsberry (2005) and the qualitative investigations by Kessler (2000), Emmer (2001), Suchan and Brewer (2000), and Slocum et al. (2004) has resulted in the considerable methodological advance of animated map design. Such work has also demonstrated that the perception of animated maps differs from that of static maps. According to Hanewinkel and Tzschaschel (2005), the variables of shape and colour are processed by the brain in a different way than movement. The first perceived group concerns direction, speed and final shape. The second level of perception constitutes moving shape and colour, and the third concerns shape of the background and initial shape. The fourth level of perception concerns the third dimension. It should be stressed that perception of these variables and elements changes with user age. According to McCloud (1993), in the case of animations we employ memory deduction, as opposed to spatial deduction in the case of static maps. Memorising the information presented in cartographic animation requires the use of both short-term and long-term memory (Sweller and Chandler, 1994). The former is employed to register the information until it is decoded, processed into a mental image and interpreted by the user (MacEachren et al., 1999). Short-term memory can be expected to be impaired with age. According to Oed (1989) the process of memorising spatio-temporal information is carried out in three stages – definition of a reference system, selective identification of the objects, and recognition of the objects and processes. Block et al (1999) have demonstrated that during the process of interpretation and analysis of spatio-temporal processes the user must perform many more logical operations than in the case of the same process concerning a spatial system. The effectiveness of this process hinges on the level of complexity (and possible excess) of spatio-temporal information, its arrangement, and the user's experience and abilities. To reduce the number of logical processes it is necessary to avoid excessive information transfer, which depends (to a considerable extent) on the correct use of visual variables.

At present, most animated maps are prepared in a single version, intended for all users. To avoid impairment of the process of spatio-temporal information transfer, especially in the case of elderly persons, it is necessary to avoid an excess of information, as well as to choose variables (and related methods of cartographic presentation) that are particularly legible for users. To make the first task easier it is possible to employ guidelines, or rules, derived from research on animated map design (including the *entities–cartotrophic method*[2] of designing simple anima-

[2] This method is based on the choice of dynamised entities and levels of measurement, the identification of types of changes and animation, as well as the selection and verification of dynamised visual and dynamic variables (and their combinations), and related methods of presentation. It employs the tables of rules concerning the correct use of combination of static and dynamic visual variables and sound variables, as well as the correct use of related methods. The name of the method makes reference to the terms: 'entity' and 'cartographic methodology': τροπος - tropos (Greek) is one of the synonyms of 'method', 'methodical'.

tions and the *entities–polystaymic method*[3] of designing complex animations proposed by the present author). In the case of the second task it is necessary to identify and analyse information on factors affecting the visual and aural perception of variables related to user age, generation and education level, then to propose lists of groups of users together with lists of variables that should be best adapted for each group. This should also make it possible to propose lists of related methods of cartographic presentation. These lists can be used to modify the general rules concerning the employment of variables and related methods of cartographic presentation, which can be applied in the procedures and methods of designing animated maps.

2 Objectives, Approach and Methods

The aim of this research was to investigate the possibilities for (and limitations on) the choice of static and dynamic visual variables and sound variables for designing simple and complex animated maps that are tailored for users belonging to different age groups, as well as to propose guidelines concerning the selection and use of these variables for users in these groups. To achieve this goal it was necessary to identify the perceptual determinants influencing the possibilities for (and limitations on) the use of particular variables in the presentation of spatio-temporal phenomena, verifying which of them can be parameterised. The next stage involved the analysis and synthesis of available information on factors affecting the visual and aural perception of variables, related to user age, generation and education level. On this basis, the main age groups of potential users were determined, the variables were ranked in terms of utility, and guidelines were proposed concerning the use of static and dynamic visual and sound variables, together with related methods of cartographic presentation for each age group. This made it possible to propose general recommendations concerning the design of animations, stemming from investigations into their perception, as well as remarks about the user interface. The utility of these guidelines was then tested on groups of representatives, employing prepared test animations. The results of these tests allowed the ranks of variables and related methods of cartographic presentation to be further refined. Based on these modified guidelines, tables of rules were proposed for the use of variables and methods for each age group, which were integrated into the author's *entities–cartotrophic method* of designing simple animations and *entities–polystaymic method* of designing complex animations. The modified animations

[3] This method is a modification of the *entities-cartrotrophic method*. It also includes the results of research on rules concerning the relationships between the combinations of variables and combinations of the methods of presentation employed in subanimations of complex/multilevel animations. The name of the method makes reference to the terms: 'entity' and to πολυσταυμη – polystaimi (Greek) - one of the synonyms of 'multilevel'.

were tested a second time, and the results were evaluated using the method of average quality metrics (Kolman, 1983; Dukaczewski, 1978).

3 Results

According to Dukaczewski (2003), it is possible to distinguish two main groups of perceptual determinants influencing the possibilities for (and limitations on) the use of variables for presenting spatio-temporal phenomena:

- determinants related to the properties of variables – the influence of: 'proper' visual variables (size, value, colour, form) and 'proper' dynamic variables (moment, duration, frequency, order, rate of change, and way of transition); 'expressive' variables (duration, order) and 'distinctive' variables (duration, frequency, rate of change); 'well dynamized static variables' (shape, size, value, colour); disruptions related to perception of variables (size versus colour)
- determinants related to external factors, concerning:
  - the ability of users to read an animated map;
  - differences in animated map perception related to the way spatial data is read;
  - differences in perception, ordering and generalisation of information related to cultural factors;
  - the individual psychosomatic properties of the user;
  - ontogenesis-related differences in perception of visual static and dynamic variables and sound variables;
  - the needs of the user, related to the type of information, its scope and level of detail.

The first group of determinants was investigated during work on the *entities–cartotrophic method* (Dukaczewski, 2005) and *entities–polystaymic method* (Dukaczewski, 2007), and results were included in the proposed tables of rules.

Determinants in the second group are of a more individual nature and their impact on the evaluation of animations is one of the main reasons for discrepancies in the results of surveys carried out about the perception of spatio-temporal information. It should be emphasised that only two determinants – the first and the fifth – can be parameterised, employing the available medical literature, opinions and statistics, as well as the results of cartographic research.

Based on the medical literature, the opinions of specialists from Warsaw Medical University, and medical statistics, it is possible to distinguish eight age groups of differing types of perception: nursery (4 – 6 years), late childhood (7 – 9), adolescence (10 – 15), advanced adolescence (16 – 19), early adulthood (20 – 30), middle adulthood (31 – 60), late adulthood (61 – 80) and advanced age (over 80)[4].

The ability of users to read an animated map depends on their age, generation and education. It should be stressed that the senses of sight and hearing are adapted to reading animation as early as nursery age. Some limitations on the ability of the sense of sight to cope with rapidly changing stimuli occur in the age of adolescence, together with mental overload as a result of intensive education. Nevertheless the perception of size, value and colour remains very good, while that of granularity and orientation remains acceptable until the end of early adulthood. A weakening in the sense of sight appears in middle adulthood and is associated with ageing and thickening of the lenses, whereby successive layers are not removed but pushed inside, thus resulting in hypermetropia. The perception of size and shape can be hampered. In late adulthood, progressive muscular atrophy decreases the ability to focus light rays on the retina. Vision may become blurred; such a user needs a much larger amount of light. At the same time, the cornea is flattened, which results in a change of the field of vision. Muscular atrophy decreases the flexibility of the iris. Response to changes in value and colour is much slower. At the same time the perception of colour is changing. The ability to distinguish smooth light waves of a length which corresponds to green, blue and violet disappears. Yellow may appear brighter, while blue is perceived as darker. It becomes very difficult to distinguish between blue and purple. The development of myopia in advanced age can complicate the perception of most visual variables. Blinks and flashes can disrupt the entire field of view, and from their disappearance the sense of sight needs a considerable time to return to normal perception.

The sense of hearing is very well prepared to receive the acoustic variables used in animations from nursery age until late childhood. Deterioration can appear at the end of adolescence, as a result of progressive corruption of hair cells in the organ of Corti. Their disappearance causes a loss of perception of sounds, ranging from high to medium. In early adulthood there begin to occur more serious changes in the functioning of the incudomalleolar articulation of the middle ear. Their calcification during an average adulthood causes decreased efficiency of the transmission of vibrations from the eardrum to the inner ear. The cerumen glands become partly desiccated. In late adulthood, the eardrum becomes thinner and frail and its muscles gradually disappear, causing it to become much harder to vibrate. In advanced age the declining number of neurons can reduce the numbers of stimuli supplied to the brain.

Conditions arising from membership of a particular generation favour individuals up to the age of 20 years (a generation with classes in computer science at the stage of primary education), as well as people in the age of early adulthood (the so-called ‘Nintendo’ generation).

Education-related conditions make people between the ages of 20 to 80 years best prepared for the perception of animated maps.

---

[4] The particular age ranges of the groups may differ depending on the country

An attempt to synthesise and rank these conditions is given in Table 5.1. In the case of the development of the senses of sight and hearing, the attributed ranks are related to medical statistics (8 means that about 80% of population have very well developed senses). In the case of education the rank is related to the percentage with good map literacy, while the generation rank reflects the influence of computer literacy.

Table 5.1 An attempt to synthesise and rank the ability of users to read animated maps related to age, generation and education

Age group (years)	Ability of users to read the animated maps related to:			
	development of senses		generation	education
	sight	hearing		
[1] nursery age (4 – 6)	8 - 9	8 - 9	8 - 9	0 - 2
[2] late childhood (7 – 9)	8 - 9	8 - 9	8 - 9	3 - 5
[3] adolescence (10 – 15)	8 - 7	7 - 6	8 - 9	5 - 6
[4] advanced adolescence (16 – 19)	8 - 7	6 - 5	8 - 9	6 - 7
[5] early adulthood (20 – 30)	8 - 7	6 - 5	6 - 4	7 - 8
[6] middle adulthood (31 – 60)	6 - 5	5 - 4	5 - 4	7 - 8
[7] late adulthood (61 – 80)	5 - 4	4 - 3	4 – 3	7 - 8
[8] advanced age (over 80)	2 - 1	2 - 1	3 - 2	6 - 7

Drawing on the medical literature, the opinions of specialists from Warsaw Medical University and medical statistics, an attempt was made to rank the perception of visual variables for users of different age groups (shown in Table 5.2).

Table 5.2 An attempt at ranking the perception of visual variables for users of different age groups

Age group (years)	1 size	2 form	3 value	4 colour	5 grain	6 orientation	7 brilliance	8 transparency
[1] nursery age (4 – 6)	8-9	8-9	8-9	8-9	4-3	4-3	5-4	7-8
[2] late childhood (7 – 9)	8-9	8-9	8-9	8-9	5-8	5-8	5-4	8-9
[3] adolescence (10 – 15)	8-9	8-9	7-8	8-9	5-8	5-8	8-9	7-8
[4] advanced adolescence (16 – 19)	8-9	8-9	7-8	8-9	7-8	7-8	8-7	7-8
[5] early adulthood (20 – 30)	8-9	8-9	7-8	8-9	7-8	7-8	7-6	7-8
[6] middle adulthood (31 – 60)	6-7	6-7	4-3	5-6	5-8	5-8	6-5	4-3
[7] late adulthood (61 – 80)	6-5	6-5	2	4-3	2	2	4-3	2
[8] advanced age (over 80)	2	2	2	2	1	1	1	2

Based on data from the same sources and the reclassification and reinterpretation of data from medical tests, an attempt was made to rank the perception of sound variables (Table 5.3).

Table 5.3 An attempt at ranking the perception of sound variables for users of different age groups (minimal values)

Age group (years)	② pitch	③ register	④ loudness	⑤ timbre	⑥ duration	⑦ rhythm	⑧ rate of change	⑨ order	⑩ frequency
[1] nursery age (4 – 6)	8	8	8	8	8	8	8	8	8
[2] late childhood (7 – 9)	8	8	8	8	8	8	8	8	8
[3] adolescence (10 – 15)	7	7	8	8	6	8	6	6	6
[4] advanced adolescence (16 – 19)	7	7	8	8	5	8	5	5	5
[5] early adulthood (20 – 30)	7	7	8	8	4	8	4	4	4
[6] middle adulthood (31 – 60)	3	3	6	6	3	6	3	3	3
[7] late adulthood (61 – 80)	2	2	5	4	2	5	2	2	2
[8] advanced age (over 80)	2	2	3	4	2	3	2	2	2

This analysis demonstrated a considerable degree of differentiation between the different age groups in terms of their abilities to perceive visual and sound variables. This data, together with the research results reported by Fabrikant and Goldsberry (2005), Koussoulakou and Kraak (1992), Harrower et al (1998, 2000) were employed to propose a set of guidelines concerning the use of static and dynamic visual and sound variables, as well as related methods of cartographic presentation for each age group, including:

- general recommendations arising from the conditions of perception;
- detailed instructions and rules regarding the use of:
  - visual static variables;
  - sound variables;
  - methods of cartographic presentation;
  - types of animation;
  - functionalities;
- recommendations concerning:
  - user interfaces.

These guidelines are available in Appendix 2 (column labelled 'A') at http://www.igik.edu.pl/upload/File/dr-dariusz-dukaczewski/_92_Appendix2.pdf. These recommendations and rules were then tested on representatives of the eight age groups. The first three groups consisted of 32 six-year-old, 86 nine-year-old, and 32 twelve-year-old children from Primary School No. 212 in Warsaw. The fourth age group was represented by 32 eighteen-year-old pupils from the 44th Secondary School in Warsaw, and the fifth group by 31 students aged 22-23 studying in different faculties at Warsaw University. Tests concerning the sixth group (31 people) and seventh group (32 people) were carried out on a population of members of ActiFrance[5]. The tests were also performed on a small population of 11 people aged over 80 years. The results of these tests may be representative of (and limited to) the population of Warsaw (in the case of the fifth, sixth and seventh group – for the population with higher education). The tested recommendations were employed to prepare eight versions of 14 simple and complex animated maps, employing different combinations of visual, dynamic and sound variables (Appendix 1). All simple animated maps were designed employing the modified *entities–cartotrophic method* (Appendix 5 at http://www.igik.edu.pl/upload/File/dr-dariusz-dukaczewski/_92_Appendix5.pdf) and modified *entities–polystaymic method* (http://www.igik.edu.pl/upload/File/dr-dariusz-dukaczewski/_92_Appendix6.pdf). The basic versions of these animations were modified, according to the guidelines concerning the use of static and dynamic visual and sound variables, as well as related methods of cartographic presentation for each age group (Appendix 2). To avoid the effect of lecture of time stages instead of animations the functionality of animations was limited to 'start', 'stop'and 'sound' buttons.

The experiments were carried out individually, face-to-face. The participants read the animations in an order ranging from simple (employing a limited number of variables) to more complex and described the message perceived. On this basis, the percentage of correct perceptions of the information presented was calculated and the degree of correctness of the applied variables and cartographic presentation method was estimated. The new ranks of the variables correspond to the decimal percentages of variables correctly perceived. In the case of static visual variables it was possible reclassify all of them individually and to test the perception of the proposed variable of aura (Dukaczewski, 2007). Due to the simultaneous use of 'proper' dynamic variables (display date, duration, frequency, order, rate of change, way of transition) assessment was carried out for six variables *en bloc*, on the basis of the available literature. The tested methods of cartographic presentation were judged to be appropriate for the age group of potential users if the percentage of correct perceptions was at least 80%. The results of this investigation are shown in Appendix 2 (column labelled 'B') at http://www.igik.edu.pl/upload/File/dr-dariusz-dukaczewski/_92_Appendix2.pdf**.**
The revised guidelines were used to modify 'the table of rules of correct combined applications of static (dynamized) visual variables and sound variables', employed in the *entities–cartotrophic method* for the design of simple animations (Dukaczewski, 2005) and the *entities–polystaymic method* of designing complex animations (Dukaczewski, 2007). This made it possible to propose five tables of rules for first, second, third-to-sixth, seventh, as well as eighth age groups. These tables

[5] interdisciplinary club of Polish French-speaking specialists, presided over by the author.

are presented jointly in Appendix 3 (at http://www.igik.edu.pl/upload/File/dr-dariusz-dukaczewski/_92_Appendix3.pdf). Similarly, the results were used to modify ‘the table of semiotic evaluation of combined applications of visual static and dynamic variables, sound variables and related methods of presentation’, yielding five tables of rules, adapted for the first, second, third-to-sixth, seventh, as well as eighth age group. These tables are presented jointly in Appendix 4 (at http://www.igik.edu.pl/upload/File/dr-dariusz-dukaczewski/_92_Appendix4.pdf). All these tables of rules can be employed in author’s *entities–cartotrophic method* and *entities–polystaymic method,* facilitating the design of animated maps for users of different age groups.

The last stage of investigation involved testing the results using the method of average quality metrics (Kolman, 1983; Dukaczewski, 1978). This method, based on the vector theory of states and qualitonomy, is employed to quantitatively assess quality on the basis of characteristics which are unmeasurable or are difficult to measure. Four of the previously tested animated maps were rebuilt according to the modified rules and adjusted to the needs of users in the second, third-to-sixth and seventh age groups. Both versions of the animated maps (‘standard’ and ‘user group adapted’) were then presented to representatives of the appropriate age groups. Most of these were the same individuals as in the first stage of tests. The task of the respondents was to assess both versions of each animation based on twelve criteria: legibility, intelligibility, visual balance, completeness/exhaustivity, order of reading, consistency, selection of signs, level of understanding the map and level of understanding the process. The answers were rated on a ten-point scale (1-10). The final results are shown in Table 5.4.

Table 5.4 The synthetic (average) indicator values of assessment of ‘standard’ (A) and ‘user’s group adapted (B) animated maps

Animations	Age groups		
	[2]	[3], [4], [5], [6]	[7]
I A	0.7033	0.7000	0.6500
I B	0.8560	0.8700	0.8700
IIA	0.6400	0.6000	0.6200
II B	0.8800	0.8860	0.8860
III A	0.7130	0.7060	0.6900
III B	0.8900	0.8900	0.8800
IV A	0.6300	0.626	0.6000
IV B	0.8800	0.8800	0.8700

These results demonstrate that the animated maps rebuilt according to the proposed modified rules received better evaluations than their ‘standard’ versions. This confirms the usefulness of the solutions developed. Although the differences in value are admittedly not large, it should be borne in mind that even the ‘standard' versions of the animated maps had been prepared correctly from the semiological point of view.

Conclusion

The aim of this research was to contribute to the discussion about the possibilities for (and limitations on) the design of animated maps that are better adapted to the needs of users belonging to different age groups, through the application of solutions based on the choice of variables specifically tailored to a given user age group. Another objective was to further advance the cartographic animation methodology itself. Analysis of the available literature, consultations with medical specialists, a review of medical statistics and specially organised tests have demonstrated the need to diversify the versions of animated maps for users of different age groups. Moreover, the research has made it possible to formulate recommendations for the design of animations. A review of available research on perception enabled general and detailed rules to be defined, and a table of guidelines to be devised regarding the use of visual and dynamic variables, methods of presentation, types of animation, and interface functionality. This data was used to refine the tables of correct use of variables, as well as methods and related combinations of variables. This then resulted in the preparation of separate versions of these tables, taking into consideration the specific perception of variables and methods by users of different age groups. These tables of rules can be used to design animations tailored to specific age groups, employing the *entities–cartotrophic* and *entities–polystaymic* methods.

The research has also demonstrated a few problems related to the availability of medical data (especially medical statistics, which are partially classified information). The results presented here are (thus far) representative and valid for the tested population in Warsaw. However, it should be stressed that the same or a similar procedure can be carried out for other populations.

Finally, the evaluation that was carried out showed that the proposed solutions are indeed operational and may serve as a useful aid for the design of animated maps.

References

Bétrancourt M, Morrison JB, Tversky B (2002) Animation: can it facilitate? Interationational Journal of Human – Computer Studies 57:247-262

Block C, Köbben B, Cheng T, Kuterema A (1999) Visualization of relationships between spatial patterns in time by cartographic animation. Cartography and Geographic Information Science. 26(2):139-151

Dukaczewski D (2003) Metody prezentacji dynamiki za pomocą zmiennych graficznych możliwych do stosowania w elektronicznych animacjach kartograficznych. IGiK – Instytut Geodezji i Kartografii, Warsaw

Dukaczewski D (2005) Entities – cartotropic method of selection of static and dynamic variables for temporal cartographic animations. In: Proceedings of the XXII rd ACI/ICA International Cartographic Conference 2005, Sociedad Española de Cartografía, Fotogrametría y Teledetección – SECFT, Instituto Geográfico Nacional, A Coruña, 9 - 16 July 2005

Dukaczewski D (2007) Method of choice of variables and cartographic presentation methods for complex cartographic animations. In: Proceedings of the XXIII International Cartographic Conference ACI/ICA 2007, Roskart, Moscow, 4 - 10 August 2007

Dukaczewski H (1978) Badanie wyrobów metodą uśrednionych znamion jakości, Ośrodek Doskonalenia Kadr Ministerstwa Przemysłu Chemicznego, Police

Emmer NM (2001) Determining the effectiveness of animations to represent geo-spatial temporal data: a first approach. In: 4th AGILE Conference on Geographic Information Science 'GI in Europe: Integrative, Interoperable, Interactive' Brno, 19 – 21 April 2001

Fabrikant S, Goldsberry K (2005) Thematic relevance and perceptual salience of dynamic geovisualisation displays. In: Proceedings of the XXII rd ACI/ICA International Cartographic Conference 2005, Sociedad Española de Cartografía, Fotogrametría y Teledetección – SECFT, Instituto Geográfico Nacional, A Coruña, 9 - 16 July 2005

Hanewinkel C, Tzschaschel S (2005) Teaching with animated maps - the use of the German National Atlas for schools. In: Proceeding of the Joint ICA Commissions Seminar on Internet-based Cartographic Teaching and Learning 'Atlases, Map Use, and Visual Analytics' Madrid, 6-8 July 2005

Harrower M, Griffin AL, MacEachren A (1998) Temporal focusing and Temporal Brushing: Assessing their Impact in Geographic Visualization. In: Proceedings of the 7th Annual Conference of Polish Spatial Information Association, Warsaw

Harrower M, MacEachren A, Griffin AL (2000) Developing a geographic visualization tool to support earth science learning. Cartography and Geographic Information Science 27(4):279-293

Kessler FC (2000) Focus groups a means of qualitatively assessing the U-boat Narrative, Cartographica 37(4):33-60

Kolman R (1983) Badanie wyrobów metodą uśrednionych znamion jakości, Towarzystwo Naukowe Organizacji i Kierownictwa, Bydgoszcz

Koussoulakou A, Kraak MJ (1992) Spatio – temporal maps and cartographic communication. The Cartographic Journal 29(2):101-108

Lowe RK (1999) Extracting information from an animation during complex visual learning. European Journal of Psychology of Education 14(2):225-244

Lowe RK (2003) Animation and learning: selective processing of information in dynamic graphics. Learning and Instruction 13:157-176

Lowe RK (2004) Interrogation of a dynamic visualization during learning. Learning and Instruction 14:257-274

MacEachren AM, Wachowicz M, Edsall R, Haug D, Masters R (1999) Constructing knowledge from multivariate spatiotemporal data. The International Journal of Geographic Information Science 13(4):311-334

McCloud S (1993) Understanding comics. Kitchensink Press, Northampton

Midtbø T (2003) Multivariable 'seeing maps' through interactive animations. In: Proceedings of the ACI – ICA International Cartographic Conference, Durban, 10 – 16 August 2003

Oed F (1989) Syntetische Welten – Visuelle Kommunikation mit modernen Medien. In: Stankowski A, Duschek K (ed.) Visuelle Kommunikation, Berlin

Opach T (2005) Semantic and Pragmatic Aspect of the Transmitting Information by Animated Maps. In: Proceedings of the XXII rd ACI/ICA International Cartographic Conference 2005, Sociedad Española de Cartografía, Fotogrametría y Teledetección – SECFT, Instituto Geográfico Nacional, A Coruña, 9 - 16 July 2005

Peterson MP (1998) That interactive thing you do. Cartographic Perspectives 29:3-4

Schnotz W, Böckheler J, Grzondziel H, Gärtner I, Wächter M (1998) Individuelles und kooperatives Lernen mit interaktiven animierten Bildern.Individual and co-operative learning with interactive animated pictures. Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology, Hans Huber AG, Bern 1998 (2/3):135-145

Slocum TA, Block C, Jiang B, Koussaukolou A, Montello DR, Fuhrmann S, Hedley NR (2001) Cognitive and usability issues in geovisualization. Cartography and Geographic Information Science 28(1):61–75

Slocum TA, Sluter RS, Kessler FC, Yoder SC (2004) A qualitative evaluation of Map Time, a program for exploring spatiotemporal point data. Cartographica 39(3):43-68

Suchan TA, Brewer CA (2000) Qualitative methods for research on mapmaking and map use. The Professional Geographer 52:145-154

Sweller J, Chandler P (1994) Why some material is difficult to learn. Cognition and Instruction 12 (3):185-233

Appendix 5.1. List of basic versions of animated maps employed for performed tests

Simple synthetic animations	**Complex synthetic animations**
monomodule, monolevel	monomodule, monolevel
Melting of Jakobshavn glacier 1850 – 1964 (MCa, 1, 2, 3, 4, 7; ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩) Land use changes in Izerskie mountains 1767 – 1994 (MCa, Sβa, 1, 2, 3, 4, 7, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩) Europe. Dynamics of gross national products 1995 – 2001 (Kγc, 1, 2, 3, 4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩) Employment in industry of Warsaw 1913 – 1995 (Kdαc, 1,2,3,4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)	Activities of world space centers 1980 – 1990 (Kdαc/ Sαa 1, 2, 3, 4, 7, 8, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)
	monomodule, multilevel
multimodule, monolevel	Number of operations in Polish airports 1992 – 2003 (Sαa / Sαa, 1, 2, 3, 4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)
Custom offices in Bogatynia region 1990 – 2003 (Sαa, 1, 2, 3, 4, 7, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)	
multimodule, multilevel	
Operations of Heli Air Monaco (Sβa 1, 2, 3, 4, 8, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)	
Simple analitical animations	**Complex analitical animations**
monomodule, monolevel	monomodule, monolevel
Town charters in Wielkopolska region X – XX century period (Sαb 1, 2, 3, 4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)	Dynamics of employment in Warsaw steel mill 1973 – 1995 (Kdαc/Kdαb, 1, 2, 3, 4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)
	monomodule, multilevel
	Europe. Part of the gross national product destined for the higher education and number of higher study diplomas (per 100 persons) in the 1997 – 2001 period (Kdαc/Kdαc, 1, 2, 3, 4, 7, 8, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)
multimodule, monolevel	
Trans-shipments in Polish ports 1946 – 1994 (Kdαc, 1, 2, 3, 4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)	
multimodule, multilevel	multimodule, multilevel
Air connections of West Berlin (1976 – 1990) (Kdβc 1, 2, 3, 4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)	Tourist infrastructure in Rybi Potok Valley (Tatra Mountains) 1836 – 2000 (Sαa/Sβa 1, 2, 3, 4, ②, ③, ④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩)

Key:

Static visual variables

1 size; 2 form; 3 value; 4 colour; 5 grain; 6 orientation; 7 brilliance; 8 transparency; 9 aura; dynamised variables are underlined

Sound variables

② Pitch; ③ Register; ④ Loudness; ⑤ Timbre; ⑥ Duration;⑦ Rhytm; ⑧ Rate of change; ⑨ Order; ⑩ Frequency

Entities / measurement levels

α point entities; β line entities; γ area entities;

a - nominal scale measurable entities; b - ordinary scale measurable entities; c - quantitative scale measurable entities

Methods of presentations

Sαb – Ordinary point signatures,

Kαb – Ordinary point choropleth maps,

Kdαb – Ordinary point cartodiagrams

Kc – Dot method,

Sαc – Quantitative point signatures,

Kαc (cs) – Quantitative point choropleth maps,

Kdαc (cs) - Quantitative point cartodiagrams,

Sβb – Ordinary line signatures,

Kβb – Ordinary line choropleth maps,

Kdβb - Ordinary line cartodiagrams

Ic – isoline maps,

Sβa – Qualitative line signatures,

Sβc – Quantitative line signatures,

Kβc (cs) - Quantitative line choropleth maps,

Kdβc (cs) - Quantitative line cartodiagrams

KαBb – Ordinary Bertin’s choropleth map,

Kγb – Ordinary area choropleth maps,

KDγb – Ordinary dasimetric choropleth maps,

KαBc - Quantitative Bertin’s choropleth map,

Kγc(cs) - Quantitative area choropleth maps,

KDγc(cs) - Quantitative dasimetric choropleth maps,

Sαa – Qualitative point signatures

MCa – Chorochromatic method maps,

MZa – Range maps

Optimising the Selection of a Number of Choropleth Map Classes

Ali M. Al-Ghamdi[1]

Abstract A simple yet a valid question often asked by GIS users is: what is the "optimal" number of choropleth map classes for a given data set? This question is barely addressed in the literature, however. In this present work, a method is therefore proposed and named the “Weighted Number of Classes Index (WNCI).” It proposes an optimisation approach according to which the resultant class break values are weighted based on the numbers of their occurrences within a set of classification runs. This is followed by normalising the total weight for each classification run; the classification run that has the highest total normalised weight is chosen, and its associated number of classes is considered comparatively the best option for the given data set. Using seven data sets, the results showed that the WNCI method rationale and performance appeared to be unaffected by the types of data sets used, because producing the highest WNCI value for each dataset is possible, regardless of the dataset values and distribution. Further enhancement of the WNCI was proposed from a digital implementation perspective.

1 Introduction

Because of their popularity, choropleth maps have attracted considerable attention in the cartographic and geographic research communities. However, revisiting this type of mapping process and form is necessary, principally because of its integration within geographic information systems (GIS). Cartographically, the challenge exists partly in properly selecting both the method of classification and the number of classes. If these two tasks are successfully realised, the geographical pattern inherent in the data will be exposed, and map communication would be enhanced. Researchers have emphasised that the number of classes and the selected classification method greatly influence a map-reader’s ability to read and interpret a

A. M. Al-Ghamdi
King Saud University
Department of Geography
Riyadh, Saudi Arabia
e-mail: aghamdi1@ksu.edu.sa
Tel: +966503102013

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_6,

choropleth map (Armstrong et al. 2003). The ultimate goal of any thematic map is to convey a valid impression of the spatial data patterns.

Most users are exposed to and limited by the available mapping options. In current mapping technologies, the user must usually experiment with various tests, and then select the classification results that suit certain criteria or needs. However, given that there are almost unlimited options, users with less training find it difficult to determine the most appropriate number of classes. This process is now partially supported by specific customised software; by using various visualisation tools and options. Nevertheless, only a trained user is capable of reaching an optimal decision using such software. It is therefore desirable that there is a better choice for both types of user; an automatic method capable of making this decision for them.

This study endorses the idea that a multitude of factors or considerations affect choropleth map production. Because these factors are barely separable, numerous issues arise and require thorough investigation, preferably from different perspectives. This study focusses only on the issue of selecting the number of classes. The approach adopted in this study is to examine this issue from an optimisation perspective, and observe how and to what extent certain factors may affect its validity. This study believes that optimising class number selection by using a certain weighting process is a solution to determining how many classes should be used (provided users are content with their selected classification schemes). An optimisation methodology is proposed and examined with reference to data sizes and types of data distribution. The optimisation process considered two main cartographic considerations; the preferred range of class numbers and "less is more."

2 Background

Numerous researchers have provided detailed overviews of the range of methods commonly employed by cartographers as well as the necessary computational algorithms. Statistical approaches to data classification are still recommended (MacEachren 1995, Robinson et al. 1995, Dent 1999, Slocum 1999). One of the most adopted and cited statistical classification methods is the Jenks optimal algorithm (Jenks 1977). The Jenks optimal algorithm has been proven optimal for data classification because it is based on quantifiable homogeneity and cluster concepts in statistics (Smith 1986, North 2009). The method was originally presented by Jenks (1977), and is based on Fisher's (1958) article 'On Grouping for Maximum Homogeneity'. This method minimises classification error and variation within classes that are measured as the sum of absolute deviations around class means (Jenks and Caspall 1971, Jenks 1977). The Jenks algorithm is referred to as an optimal classification method, and has been implemented in certain mapping software. Whereas the programming implementation of the method might vary by software, the method concept remains the same. Some GIS and digital mapping software provide this method as the default method for data classification, and occasionally call it Natural Breaks, such as in the ArcGIS package used in this study. The Jenks method requires that the user first enter the desired number of classes; this is not a major concern for professional users or cartographers, but it is for untrained users (North 2009). Although it is reasonable to regard the natural breaks method as a standard and well-tested data classification method, its adaptation in this study neither implies nor endorses its suitability for all classification tasks.

Research into cognition and perception promotes the use of a limited number of classes because discerning differences between a limited number of shadings is difficult (Mersey 1990, Slocum et al. 2005). Encouraged by Miller (1956), cartographers commonly use numbers of classes ranging between five and nine. Miller stated that the number of objects an average person can hold in their working memory is 7 ± 2. Previous research (Miller 1956, Evans 1977, Mersey 1984, 1990, Cromley 1995) recommended five to seven classes for static maps, and other researchers (Harrower 2003) have recommended less than five classes (as few as three) for dynamic map representations. Increasing the number of classes increases the map complexity, although this depends on a user's cartographic skill and experience (Evans 1977). As a compromise, Mersey (1990) suggested that five classes would be suitable, and possibly more if a colour scheme is used. However, ambiguity remains as to how many classes are cartographically sufficient (e.g. within the three to nine range); that is, what is the number of classes below which the mapped data becomes oversimplified, and beyond which the map either begins to appear too complex or beyond which further classes no longer legibly enhance the content?

Many researchers have studied the effect of the number of classes on map characteristics such as pattern (Muller 1976), colour (Mersey 1984, 1990, Gilmartin and Shelton 1989), classification accuracy (Jenks and Caspall 1971, Coulson 1987, MacEachren 1985), data distribution (Evans 1977, Cromley and Campbell 1994, Cromley 1995), and the effect of the number of classes as well as data uncertainty on the robustness of choropleth classification schemes (Xiao et al. 2007). Others, however, have proposed another alternative: creating unclassed choropleth maps. Tobler (1973) was the first to propose this approach. Dobson (1973) contested Tobler's results, but other researchers have continued to explore this concept (e.g., Swift and Nishri 1993).

Studies most relevant to the present study have been presented by Jenks and Caspall (1971), Stegna and Csillag (1987), Cromley and Campbell (1994), MacEachren (1985), Cromley (1995), and North (2009). These researchers directly addressed the issue of how to determine the optimal number of classes. Three studies (Jenks and Caspall 1971, Coulson 1987, MacEchren 1985) proposed using accuracy indices, such as the GVF, to determine the proper number of classes. However, determining the proper number of classes based solely on classification accuracy and optimal classification methods does not necessarily solve the problem from the user perspective because the user must decide between suboptimal choices. For example, users might find it difficult to choose between two successive numbers of classes if their classification accuracy values appear to differ only by a fraction of decimal places (e.g. the difference between 0.9564 and 0.9566). Stegna and Csillag (1987) suggested that the number of classes is statistically determined based on minimising the redundancy of information content. They used iterative t-tests to determine the number of classes that are statistically significant for separability, and thus the optimal number of classes.

Cromley (1995) studied unclassed maps versus classed maps from a data distribution perspective. He suggested that a common drawback to perceptual studies is that no discussion exists about whether a single number of classes should represent every data distribution or whether the number of classes should vary with the data distribution. He stressed that the focus should be on determining the proper number of classes required to represent the data distribution, whether for an unclassed or classed choropleth map representation. He then proposed an optimal classification scheme that simultaneously determines the number of classes and the intervals between class breaks by using a hypothesised level of graphic value discrimination for data distribution.

North (2009) is more specific and relevant. He proposed using an iterative t-test methodology to identify a statistically significant number of classes through programming. North briefly presented and explained a programming pseudo-code to demonstrate his method. Although the t-test methodology has been used in data classification (e.g., Stegna and Csillag (1987)), North suggested its use within a programmable mechanism for determining the appropriate number of classes for any data set. However, North indicated that his proposed method was both a work in progress and a theoretical solution to statistically determining the optimal or appropriate number of classes.

Because of the complex nature of, and the increasing interest in, choropleth mapping and advances in GISs and cartography, choropleth map classification research explored geovisualisation through dynamic cartographic displays. This approach provides users with a method for data exploration, analysis, and visualisation by using interactive tools (Andrienko et al. 2001, Xiao & Armstrong 2006, Egbert and Slocum 1992). However, using this methodology implies that a user is potentially an expert and is interested in more than a simple cartographic output.

Reviewing the literature indicates that the emphasis was on the effects of the selection of the number of classes and the classification schemes on the resultant pattern complexity and classification accuracy. However, GIS users must determine how many classes should be used during choropleth map production tasks. In other words, what is the most appropriate number of classes that best reflects the statistical nature of a dataset, as well as considering all or some of the established cartographic principles and recommendations? Cartographic considerations might include a host of issues, such as map design, scale, map purpose, and audience. Determining the number of classes appears to be a context-dependent task, thereby increasing the difficulties of proposing specific guidelines for every possible case; however, this should encourage, not deter, further research. This chapter focuses on developing a novel straightforward methodology for determining an appropriate number of classes specifically for choropleth maps with reference to the Jenks classification scheme.

3 Methodology

3.1 Proposition

The proposed methodology first assumes that the natural breaks method (based on the Jenks optimal method) is used. If we consider that the user is experimenting with a limited number of classes (from five to nine, as based on the literature), the user would then produce a set of five classification runs that correspond to five, six, seven, eight, and nine classes. The user learns that each class break value occurs either once, twice, three, four, or five times. The repetition (occurrence) of a class break value more than once implies that this value appears to be relatively important or significant, and this significance increases as the repetition increases. To quantify this repetition (importance) for each class break value, some type of weight (e.g. a numerical value) should be assigned, and for this task the number of repetitions itself is adequate. By adding these weight values together for each classification run, the total weight value is then regarded as being indicative of the relative importance or optimality of the respective classification run. From a cartographic perspective, a final step is required: this weight should be refined, referred to here as 'normalised'. This ensures that a classification run with nine classes is not necessarily selected because the weight values in this level might become the highest, subsequent to the addition process. The cartographic consideration adopted here for normalisation is based on a cartographic principle: the best map scale is the smallest one that can accommodate the map details (Robinson et al. 1995). Adoption of this principle is further endorsed by the conclusion reached in Cromley's work (1995). Accordingly, this principle was translated to be the least number of classes with the highest weight, which is preferred; this is achieved by dividing the total weight of each classification run by its respective number of classes. The calculation of the entire process is therefore accomplished in three steps, with reference to a set of five classification runs, as illustrated below.

3.2 Calculation

Three consecutive steps are required: (1) calculating the individual weights of class break values in each classification run, (2) calculating the total weight of each classification run, and (3) normalising the total weight for each classification run.

In the first step, each class break value is assigned a weight according to its occurrence in each classification run. This weight is initially assigned a value of “1,” corresponding to one occurrence; the maximum weight a class break value assumes is “5” if the class break value occurred in all the five classification runs. The final (total) weight for each class break is the result of how many times this class break value occurred within the five classification runs. Because the first and last values are the two limiting values of the data set, they are not regarded as class breaks per se; consequently, they are omitted. These two value limits serve the purpose of labelling the map legend.

Once the total weight for each class break value is known in the first step, the second step is to calculate the total weight for each classification run. In this step, the individual weights of the class break values in each classification run are added. This total weight is considered to reflect the weight of the number of classes corresponding to its respective classification run. The total weight for a given classification run (number of classes) is expected to differ from that of the other classification runs.

In the third step, the total weight of each classification run is divided by the number of classes, representing a final normalised value (index) as being the indicator of the final (optimised) weight of the number of classes. The number of classes with the highest normalised weight is then selected to represent the optimal number of classes.

This proposed method is called the Weighted Number of Classes Index (WNCI), and can be written mathematically as follows (Eq. 1):

$$WNCI = \sum_{i=1}^{n} w_i / n \tag{1}$$

where *WNCI* is Weighted Number of Classes Index, n is the number of classes, and w is the total number of occurrences (weight) of each class break value (i) resulting from five classification runs (from five to nine classes).

The performance of the method is examined according to two main considerations: the effects of the data-set size and data distribution. To accomplish this, seven datasets are used, which vary in size and distribution.

As to the scope of the method application, the proposed method must only be used in conjunction with the natural breaks (Jenks) method. The method can only be regarded as a selection measure that relates to the number of classes, and should not be seen as similar to, or concerned with, other classification accuracy measures. The proposed measure only provides a map user with a new decision-making tool for choropleth mapping. The proposed method is only for users who need to optimise their selections of the number of classes automatically; they may otherwise opt to experiment with other classification methods to arrive at the solution that suits their specific needs.

3.3 Datasets

Seven datasets of varying sizes and distributions were used. The first and second datasets contain 40 enumerating units, which vary in size. These two datasets were derived from the 2005 census data for Al-Karj city, Saudi Arabia (Bureau of Statistics and General Information 2005). The first dataset represents the population density (persons per hectare), and the second represents the population percentage. The third, fourth, fifth, and sixth datasets represent the 1992 world population density (The World Bank 2012) and ESRI's free shape files). The seventh dataset represents the 2010 county population density of the United States (United States Census Bureau 2012). From a cartographic perspective, most of the unit areas that appear indiscernible at the display scale (A4 size) were deleted, for both the world and US datasets. Thus, for the world dataset, only 163 of over 200 units (countries) were selected for analysis and display, and for the US dataset, 3109 counties were selected from 3221. The summary statistics of the datasets are shown in Table 6.1.

For convenience and analysis control purposes, the third dataset represents the actual world population density in 1992 (persons per kilometre squared); the other three world datasets (the fourth, fifth, and sixth) were derived from this (third) dataset. This included artificially adjusting (changing) the original dataset (skewed distribution) to reflect different types of data distribution. Consequently, three types of data distributions were produced: normal (fourth data set), bimodal (fifth data set), and uniform (sixth data set). For convenience, especially from a cartographic design perspective, the data values of all datasets were approximated (no decimal values) except for the second dataset where the values were approximated to only two decimal places because almost half its values were below 1%. This step was performed before applying the classification process.

The seven datasets constitute an analytical framework composed of various numbers and sizes of enumeration areas, data distributions, and values. This should provide a reasonable testing context for assessing the validity and process of the proposed methodology. Datasets were processed, analysed, and displayed using ArcGIS 10.2, the ESRI GIS package. The natural breaks (Jenks) method was applied to the datasets.

Table 6.1 Summary statistics of the datasets

Datasets*	Count	Min	Max	Sum	Mean	Standard Deviation	Skewness
First	40	1.55	227.07	1721.05	43.03	45.18	2.20
Second	40	0.06	15.16	100	2.5	3.59	2.18
Third	163	0.03	871.66	14043	86.15	114.35	3.26
Fourth	163	0.03	872	71050	435.89	153.92	0.00
Fifth	163	0.03	872	72913	447.32	239.21	0.007
Sixth	163	3	872	71060	435.95	257.71	0.054
Seventh	3109	0.36	18229	260954	83.94	474.59	24.632

*Types of data distribution: first (skewed), second (skewed), third (skewed), fourth (normal), fifth (bimodal), sixth (uniform), seventh (skewed)

4 RESULTS

Table 6.2 presents a summary of the test results of the proposed method using seven datasets. The table shows the weight range for the class break values resulting from five classification runs, using five classes, starting from five to nine. The weight is the accumulative number of occurrences of each class break value calculated from five classification runs. The WNCI value is the resultant value obtained from the normalisation process, and represents the final weight of the number of classes. Also shown are the number of class break values that fell below each weight, the recommended number of classes for each dataset corresponding to the highest WNCI value, types of data distribution, and number of enumeration unit areas.

For the range of weights (numbers of occurrences), the table shows that the full range (from 1 to 5) was only applicable to the first dataset. For the next six datasets, the ranges of weights, regarding levels, were less than those of the first dataset, and their ranks varied. For instance, no class break values occurred three times (Weight 3) for the second dataset. For the third dataset, no class break values occurred four times (Weight 4), which is a result similar to that obtained for the seventh dataset. However, for the fourth and fifth datasets, no class break values occurred five times (Weight 5). For the sixth dataset, only three levels of weights corresponded to the first three ranks (1, 2, and 3).

The WNCI values can be categorised in two levels: higher and lower values, which concern two groups of datasets. The first group includes the smaller datasets (first and second), and the second group includes the larger datasets (from the third to the seventh). These values (WNCIs) were higher for the smaller datasets (from 2.4 to 3.5) than those of the larger datasets (from 1 to 2.38). However, the difference between these two ranges was only 0.28 (1.1 for

Table 6.2 Summary of the test results of the proposed method (WNCI) for optimising the selection of a number of classes

Datasets	Range of class weights*	WNCI values of no. of classes					No. of class break values under each class weight**					Recommended no. of classes	No. of enum. units areas (count)
		5	6	7	8	9	1	2	3	4	5		
Frist	1,2,3,4,5	2.4	3.5	3.43	3	3	4	2	6	8	10	6	40
Second	1,2,4,5	3	3.17	3.43	3	2.89	6	2	0	12	10	7	40
Third	1,2,3,5	2.2	1.5	1.86	1.75	1.89	14	8	3	0	5	5	163
Fourth	1,2,3,4	2	2.17	1.14	2.38	2.22	7	10	9	4	0	8	163
Fifth	1,2,3,4	2	2.33	2.14	1.63	1.63	7	10	9	4	0	6	163
Sixth	1,2,3	1.8	1	1.57	1.38	1.22	15	12	3	0	0	5	163
Seventh	1,2,3,5	1.8	1.83	1.86	2	1.89	15	4	6	0	5	8	3109

*Number of occurrences, ** Number of class break values out of the 30 values resulting from five classification runs for each dataset, excluding the first and last class limit values.

the first set and 1.38 for the second set). The lowest range among all the ranges of the seven datasets was that of the seventh dataset (0.2).

The table shows that most of the class break values fell under the lowest weights (1 and 2) for the larger datasets (from the third to the seventh). Conversely, for the smaller datasets (first and second), most class break values fell under the highest weights (4 and 5).

The results show that no specific trend or consistency existed between the recommended numbers of classes, types of data distribution, and sizes of datasets. For instance, it was recommended that the first three datasets were best classed with a different number of classes, although they were of the same type of data distribution (skewed). Furthermore, although the first two datasets had a similar number of enumeration units and a similar type of data distribution, the proposed method recommended two different numbers of classes. Conversely, the third and sixth datasets were assigned the same number of classes (5), although they were different in size and the type of distribution.

Only a few examples (four) of the resultant classification maps are reported here but they pertain to the recommended number of classes, representing the third (Fig.6.1), fourth (Fig.6.2), fifth (Fig.6.3), and sixth (Fig.6.4) datasets. When a user is given a set of various classification options resulting in a series of classed maps, ascertaining the "optimal" or most appropriate map is difficult. This methodology was proposed to aid the user in this task; it removes the burden of making the decision from the user if the user accepts the underlying principles and justifications of the method. Displaying all other maps classified here was thus deemed unnecessary.

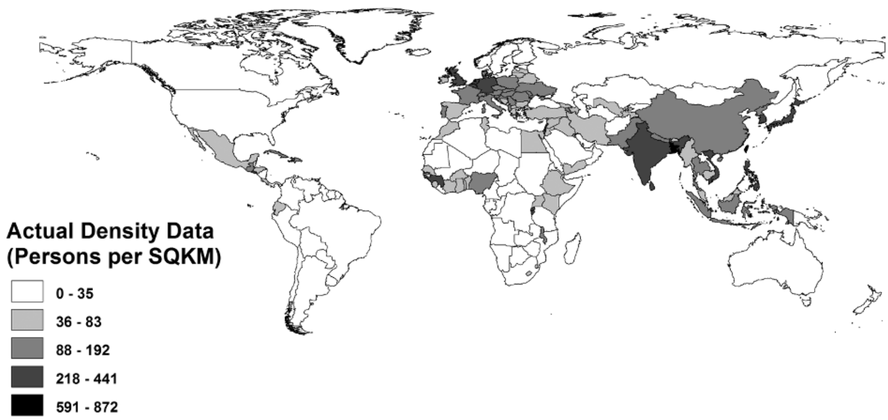


Fig. 6.1 The recommended number of classes (5) by the proposed method (WNCI) for the third dataset (World 1992 actual population density). Data distribution: skewed. Data source: The World Bank (2012).

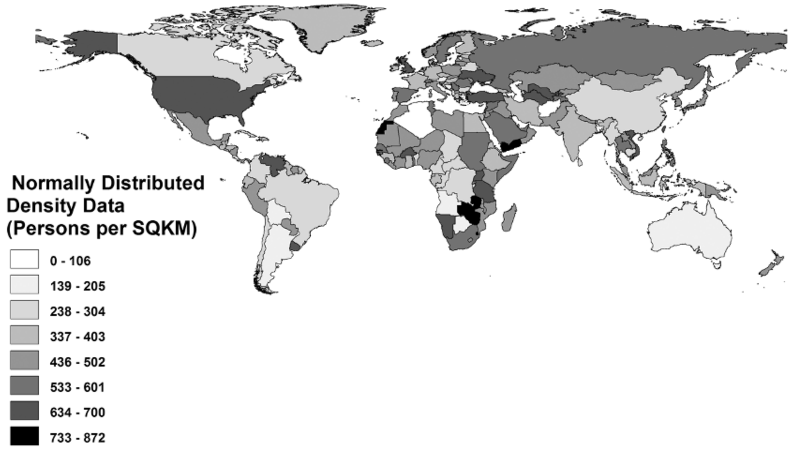


Fig. 6.2 The recommended number of classes (8) by the proposed method (WNCI) for the fourth dataset (World 1992 adjusted population density data). Data distribution: normal. Original data source: The World Bank (2012).

Fig. 6.3 The recommended number of classes (6) by the proposed method (WNCI) for the fifth dataset (World 1992 adjusted population density data). Data distribution: bimodal. Original data source: The World Bank (2012).

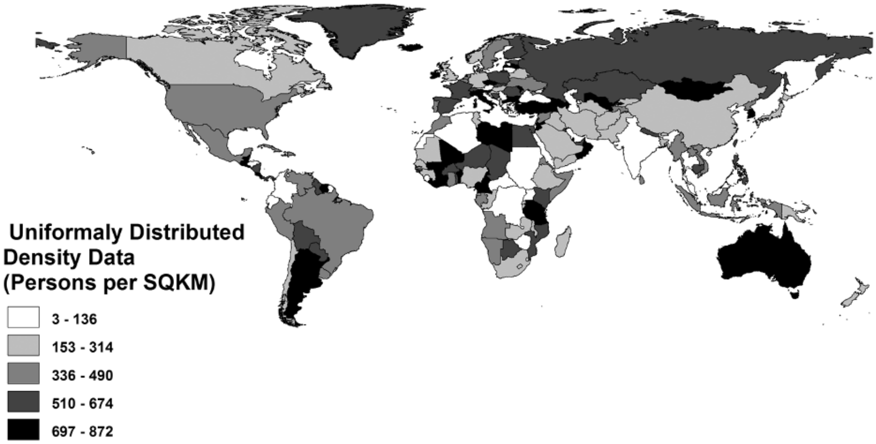


Fig. 6.4 The recommended number of classes (5) by the proposed method (WNCI) for the sixth dataset (World 1992 adjusted population density data). Data distribution: uniform. Original data source: The World Bank (2012).

5 Discussion

5.1 Effect of dataset size

The results showed that an inverse relationship exists between the size of the dataset (i.e., the number of enumeration unit areas) and the size of the WNCI values; that is, when the dataset became larger, the WNCI values became lower. This is mainly because of the effect of dataset size on the occurrence or repetition of the class break values. In other words, with a smaller dataset, the class break values tend to occur more frequently, causing most of them to fall under higher weights (i.e., 4 and 5) and, consequently, higher WNCI values. Conversely, with a larger dataset, fewer class break values occur, leading to lower weights and to lower WNCI values. However, because the WNCI values for each dataset were variable (i.e., no repetition of the same WNCI value for all classification runs occurred in a single dataset), it was possible to use the highest value as an indication of the relative significance of the respective number of classes. From this perspective, the validity of the method, regarding its ability to select the best number of classes, is not negated by the dataset size.

To account for potential effects of dataset size on the performance of the method, it is important to address the lower and upper size limits below or above which the proposed method may cease to be valid. The number of enumeration units must therefore be examined. According to the datasets used, no relationship existed between the number of enumeration units and the recommended number of classes. In other words, each dataset had its own number of classes irrespective of how large or small the number of enumeration areas was. For the lower size limit, we can infer from the results that, as long as the number of enumeration units is above the maximum number of classes (nine in this study), the proposed method is theoretically valid for application. The inference is further clarified according to a key cartographic consideration. Theoretically, the proposed method, when using a dataset with nine enumeration areas, might recommend the highest number of classes (nine). In this case, classifying nine enumeration areas using nine classes is cartographically unacceptable, or at least untenable. Furthermore, because the purpose of classification is to generalise large amounts of data into a manageable set of classes, smaller datasets, such as a dataset with nine or less enumeration areas, are more likely to be represented by five classes or less. Alternatively, a cartographer would use a different representation altogether (e.g. proportional symbols) for mapping smaller datasets.

Proposing more than nine enumeration areas for the lower dataset size limit is reasonable, below this number, the method results are invalid or uninformative. In implementing the proposed method in digital mapping or GIS software, this condition (limit) can be stated at the beginning of the programming code as follows: if the number of enumeration units is above nine, execute the method; otherwise, stop the execution. This is a practical solution because no consensus exists among cartographers regarding the definition of the lower limit.

This issue can be examined specifically for the upper limit of the dataset size by considering the results of the largest dataset (the seventh dataset). The test results showed that applying the proposed method to the seventh dataset (thousands of unit areas) did not produce noticeably different results from those of the other large dataset tests for the occurrence pattern of the class break values. The sixth dataset test resulted in lower WNCI values than the values of the seventh dataset. Because the WNCI values were variable for the seventh dataset (although with only a 0.2 difference between the lowest and largest values), it was possible to select the highest WNCI value as an indicator of the significance of the relevant number of classes. However, in an extreme case, each class break might have occurred once; consequently, the results of applying the method could be different. In this case, the total weight for the classification run of five classes is four, and its WNCI value is 0.8. By continuing the calculation, we find that the WNCI values systematically increase until they reach a value of 0.89 for the highest classification run (nine classes). Accordingly, the last classification run (nine classes) is regarded as the best option. In this specific case, the method is invalid because the recommended number of classes was a function of a systematic numerical progression of the WNCI values, which led to selecting only the largest number of classes, irrespective to the dataset characteristics. To ensure validity, such cases are easily handled programmatically. Once the method is applied, the implementation should test for the occurrence of two conditions: if all WNCI values are similar (first condition) or become increasingly larger with an increased number of classes (second condition), then the method presents a prompt "invalid results" message. The second condition can be determined by examining whether a strong positive linear relationship exists between the WNCI values and the number of classes. If the first condition occurs, the result of the method is uninformative because the method provides no solution. With such implementation, the method should address any cases related to either dataset size or any factor that is causing the method to be invalid. However, future implementations and extensive testing of the proposed method by researchers may suggest necessary amendments.

5.2 Effect of data distribution type

The results showed that similarity in the distribution among the datasets did not result in a similar number of classes, and the different data distributions did not necessarily lead to a different number of classes. For example, the first, second and third datasets all had similar distributions (skewed), but the method recommended a different number of classes for each dataset. However, the third dataset (with a skewed data distribution) and the sixth dataset (with uniformly distributed data) were both classified into five classes. Because the data range of the third dataset was different from that of the sixth dataset, the similarity in the recommended number of classes for the two datasets can be attributed to the data values and not the data distribution.

In contrast, although the results of the fourth and fifth datasets showed similarities in the weight levels and number of class break values under each weight, the recommended number of classes for each dataset was different. Because the two data distributions were different but the data values were similar, we can infer that the method appears to be sensitive to the type of data distribution, but not to the values.

In reality, a cartographer recognises that no predetermined or specific relationship exists between the type of data distribution and the number of classes. However, a cartographer does not necessarily use different numbers of classes for different types of data distribution (i.e., each case of data distribution is treated individually). Accordingly, the results of the proposed method would be less meaningful, or invalid, if different data distributions representing the same data were classified using similar numbers of classes. In other words, the performance of the proposed method regarding the selection of the number of classes appears to respond to the variability in both the data distribution and the data values. Similarities either in the type of data distribution or in the data values do not warrant using a similar number of classes for classification. The proposed method appears to conform to this mapping principle. Thus, while the proposed method maintained consistency in the calculation process, its application appeared to account for the different types of data distribution.

6 Conclusion

In this study, a novel method was developed for optimising the selection of the number of classes used in choropleth maps, the WNCI. The results of all tests showed that despite the different values, class break values, and repetition (occurrence) of class break values, the WNCI method produced consistent results according to its rationale. The proposed method did not produce similar results (number of classes) for similar dataset sizes, data values, or data distributions, and is comparable to the rationale a cartographer would follow.

The method can be implemented digitally as an add-in method within any mapping software that supports the natural breaks method. With such implementation, the user is provided with an automatic decision-making tool regarding the selection of the number of classes. Certain necessary amendments in implementation are suggested, to ensure adherence to the rationale and validity of the method from a cartographic perspective. These amendments included three application conditions. Firstly, a dataset should contain more than nine enumeration units for the lower limit of the dataset size. Secondly, no WNCI values should be similar. Thirdly, a positive linear relationship should not exist between the WNCI values and the number of classes (i.e., an increase of the WNCI values should not be a function of the number of classes). This specific conclusion endorses the idea that developing solutions for complex situations and contexts, such as the issue under discussion, requires a holistic approach which can only be achieved through powerful capabilities in programming.

The proposed method is presented for external verification. This complex and challenging research area requires the examination of various factors affecting the choice of the number of classes in the choropleth mapping process. It is recognised that further tests might lead to more general conclusions regarding the relationship between the number of classes, the dataset size, and the type of data distribution, and must be examined further by using different dataset sizes and types of data distributions.

References

Armstrong M, Xiao N, Bennett D (2003) Using genetic algorithms to create multicriteria class intervals for choropleth maps. Annals of the Association of American Geographers 93:595-623

Andrienko G, Andrienko N, Savinov A (2001) Choropleth maps: classification revisited. In: Proceedings of ICA. Beijing, China, 2:1209-1219, 6–10 August 2001

Bureau of Statistics and General Information (2005) Saudi National Census and Demographic Data. Riyadh, Saudi Arabia

Cromley RG, Campbell GM (1994) Range constrained grouping of univariate data for choroplethic displays. In: Waugh TC, Healey RG (eds) Proceedings of the Sixth International Symposium on Spatial Data Handling. Edinburgh, Scotland, 2:341 – 355, September 1994

Cromley RG (1995) Classes versus unclassed choropleth maps: A question of how many classes. Cartographica 32(4):15-27

Coulson MR (1987) In the matter of class intervals for choropleth maps: with particular reference to the work of George F. Jenks. Cartographica 24(2):16-39

Dent BD (1999) Cartography: Thematic Map Design (5th Edition). WCB/McGraw-Hill, Dubuque, IA

Dobson MW (1973) Choropleth maps without class intervals?: a comment. Geographical Analysis, 5(3):358-360

Egbert SL, Slocum TA (1992) EXPLOREMAP: an exploration system for choropleth maps. Annals of the Association of American Geographers 82(2):275-288

Evans IS (1977) The selection of class intervals. Transactions of the Institute of British Geographers, New Series 2 (1):98-124

Fisher WD (1958) On grouping for maximum homogeneity. Journal of the American Statistical Association 53:789-798

Gilmartin P, Shelton E (1989) Choropleth maps on high resolution CRTs: the effect of number of classes and hue on communication. Cartographica 26(2):40-52

Harrower M (2003) Designing effective animated maps. Cartographic Perspectives 44:63-65

Jenks GF, Caspall FC (1971) Error on choroplethic maps: definition, measurement, reduction. Annals of the Association of American Geographers 61(2):217-244

Jenks GF (1977) Optimal data classification for choropleth maps. Department of Geography Occasional Paper no. 2. Lawrence: University of Kansas

McEachren AM (1985) Accuracy of thematic maps/implications of choropleth symbolization. Cartographica 22(1):38-58

MacEachren AM (1995) How Maps Work: Representation, Visualization, and Design. Guilford Press, New York

Mersey JE (1984) The Effects of Color Scheme and Number of Classes on Choropleth Map Communication (Map Use, Design, Thematic Mapping). Dissertation, University of Wisconsin - Madison

Mersey JE (1990) Color and thematic map design: the role of color scheme and map complexity in choropleth map communication. Cartographica 27, Monograph 41, 157 pages

Miller GA (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review 63(2):343-355

Muller J-C (1976) Number of classes and choropleth pattern characteristics. The American Cartographer 3(2):169-176

North MA (2009) A method for implementing a statistically significant number of data classes in the Jenks Algorithm. Fuzzy Systems and Knowledge Discovery 2(1):202-207

Robinson AH, Morrison JL, Muehrcke PC et al (1995) Elements of Cartography. Wiley, New York

Slocum TA, McMaster RB, Kessler FC et al (2005) Thematic Cartography and Geographic Visualization. Prentice Hall, Upper Saddle River, NJ

Slocum TA (1999) Thematic Cartography and Visualization. Prentice-Hall, Upper Saddle River, NJ

Smith R (1986) Comparing traditional methods for selecting class intervals on choropleth maps. Professional Geographer 38(1):62-67

Stegna L, Csillag F (1987) Statistical determination of class intervals for maps. The Cartographic Journal 24(2):142-146

Swift M, Nishri D (1993) Using SASIGRAP software to create unclassed choropleth maps. In: Proceedings of the SAS 18 Users Group International, New York, 9 – 12 May 1993

Tobler WR (1973) Choropleth maps without class intervals. Geographical Analysis 5(3):262-265

The World Bank (2012) World Population (Total), 1992.http://data.worldbank.org/indicator/SP.POP.TOTL?order=wbapi_data_value_1992+wbapi_data_value&sort=asc&page=4. Accessed 14 May 2012

United States Census Bureau (2012) 2010 Census TIGER/Line Shapefiles, Pre-Joined with 2010 Census Demographic Profile.http://www.census.gov/geo/www/tiger/tgrshp2010/2010DP1.html. Accessed 20 October 2012

Xiao N, Armstrong MP (2006) ChoroWare: a software toolkit for choropleth map classification. Geographical Analysis 38(1):102–121

Xiao N, Calder C, Armstrong M (2007) Assessing the effect of attribute uncertainty on the robustness of choropleth map classification. International Journal of Geographical Information Science 21(2):121-144

Elaboration of the Technology for Development of Planning Freedom Maps Using GIS Tools

Krystyna Kurowska, Hubert Kryszk and Ryszard Cymerman

Abstract The rational organisation of space, even when subjected to certain principles, offers much better opportunities than the spontaneous efforts of individuals working for their own benefit only. The use of space is determined to a significant extent by the framework of planning freedom which results, among other things, from the conditions of the environment, legal regulations, existing status of land development, spatial organisation, economic principles of the economy, conditions resulting from the use and ownership structure, and social conditions. The environmental conditions and limitations to the disposal of the given space resulting from environment protection are of major importance. Those conditions limit planning freedom, creating a framework for planning solutions, and they are necessary for rational space management. The development of universal principles for creating thematic maps supporting the planning process, in this case the planning freedom map, represents a useful tool (Peng et al. 2011). For this purpose, the notion of *planning freedom restricted by environmental considerations* and the new baseline document for planning purposes – the map of planning freedom (ultimately referred to as the *map of planning freedom restricted by environmental considerations)* were developed. The opportunities for using them in planning practice are also presented.

1 Introduction

The main goal of spatial planning is to search for locations in space, for various activities (Cartwright 1973). The location of a specific planned function depends on numerous diversified factors, conditions and limitations (Alterman and Page 1973). The higher the number of limiting factors the more difficult it is to find an appropriate location and thus he more restricted is the freedom of the planner in selecting options (Gans 1970). The location of a specific activity depends on envi-

K. Kurowska (✉)
University of Warmia and Mazury, Olsztyn, Poland
e-mail: krystyna.kurowska@uwm.edu.pl

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_7,

ronmental conditions, social needs, economic potential and technical feasibility. One of the basic rules of the Polish planning system involves preservation of the requirements of sustainable development. Sustainable development is defined as the process of socioeconomic development that integrates political, economic and social measures maintaining balance in the entire world of the nature, and sustainability of basic natural processes for the purpose of satisfying the basic needs of the local communities and individuals in the current and future generations (No 62, item 627, 2001).

The necessity of preserving valuable areas of the environment forced society to protect various objects against total or partial use or destruction. This is one of the most effective instruments for preserving the environment in its natural state. Conservation as an activity protecting the environment thus represents a limitation in the use of space and the allocation of the space for specific functions (Cymerman and Kurowska 2001).

Solving the problem of the limited availability of space and the necessity of maintaining certain parts of it in the natural state should take account of the sustainable development rules. On the other hand the increasing needs (pressures) for economic use of the individual fragments of that space compels us to undertake research into the issue of the magnitude of those limitations in spatial planning that result from various forms of protection. GIS is currently used everywhere presentation of a certain issue in the spatial dimension is necessary. Modern digital map technologies offer the possibility of complete vectorisation of spatial data. Planning and spatial management are areas where GIS offers extensive opportunities for application. Currently, in Poland, GIS technologies are used first of all in the process of planning documentation development, mainly as a tool for spatial data gathering and presentation. Development of GIS technologies has meant that in recent times the system has become supportive in making planning decisions (MacEachren et al. 2005). The technological progress and changes in legal regulations that have taken place in recent years, as well as the INSPIRE Directive, stimulated numerous institutions to position spatial data on the internet (Rannestig 2009, Rodriguez et al. 2009). INSPIRE (INfrastructure for SPatialInfoRmation in Europe) is a collection of legal, organisational and technical measures with related services offering general access to spatial data within the area of the European Union. It is designed to support legislators in taking decisions and actions that could have direct or indirect influence on the environment (Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007).

In Poland, the internet portal www.geoportal.gov.pl is the major location for making the resources of digital maps available on the internet (Ažman and Petek 2009). The portal currently represents a truly innovative solution to making spatial data available and providing access to data collections in the possession of public administration (Dukaczewski and Bielecka 2009). Thanks to the Geoportal, spatial information infrastructure services are available not only to the selected units of administration responsible for state-operated registers, but also to a wide group of individuals and legal entities.

In Poland the Map of Environment Protection Restrictions is used. It is a thematic map. It presents the status of the natural environment, and the causes as well as consequences of negative and positive transformations taking place in that environment. The map mainly addresses the environment protection institutions and offices as well as decision makers and planners at the regional, voivodeship, county and commune level. It is useful in the positioning of new economic (industrial) facilities as well as municipal projects (including housing) and centres of recreation and similar spaces. A map at the scale of 1:50 000, which is not good enough represents the source cartographic material for development of the maps of environment protection restrictions in smaller scales, as well as other related thematic maps. A solution is provided by the development of a map of the planning freedom restricted by environmental considerations to meet the needs of spatial planning with particular focus on local planning.

2 Methodology

The main goal of this chapter is the development of technical assumptions for use in the development of freedom maps as useful thematic maps in supporting the process of spatial planning, based on the example of Poland. The fundamental stages in creating the planning freedom map must involve determination of the magnitude of influence of protective activities on planning activities, and identification of the scale of the problem, its extent and scope. This goal is accomplished through the following specific objectives:

1) The notion of planning freedom restricted by environmental consideration has been defined and described theoretically (the subject literature as well as my own considerations and observations were used, using the philosophical understanding of the notion of freedom as the baseline).
2) The influence of protection limitations on planning freedom was determined: an analysis of legal regulations on environment protection was conducted, which resulted in the development of the inventory of the existing forms of environment protection as well as the related type and scope of limitations in space management. The specification of the forms of protection obtained was analysed considering their importance and use in planning studies. The magnitude of the influence of different protection limitations on planning freedom was determined by means of a questionnaire-based survey.
3) Creation of a technique for the development of maps of planning freedom restricted by environmental considerations – maps of areas with restrictions to planning resulting from environment protection were elaborated (own methodology for the determination of areas with different levels of planning freedom related to limitations resulting from environment protection was elaborated. This technique considers not only the spatial forms of protected objects but also cases of the overlapping of different

forms of protection and the presence of areas with different scopes of protection.

4) The potential for using the developed map of planning freedom restricted by environmental considerations in planning studies – its usability and applications were shown (analysis of the procedure for elaboration of the study of the conditions and directions of spatial development of a commune as well as of the local spatial development plan, analysis of their findings and indication of places where protection conditions are considered and can determine allocation of the given area was conducted).
5) Creation of a sample map of planning freedom restricted by environmental considerations, as developed for the selected research object (selected units within the commune of Łyse, County of Ostrołęka, MazowieckieVoivodship.
6) Using GIS tools, a model of the map of planning freedom restricted by environmental considerations was developed. GIS software – MapInfo Professional, which works with different formats of spatial data and many types of databases – was used for specific analyses (Hopfer and MacEarchren 2007). The data imported to the software is presented in the form of layers consisting of the map and the table of attributes. This allows integration of data originating from, for example, Microsoft Excel or Access files or servers of databases (Oracle, etc.), with the map and then their visualisation.

3 Definition and theoretical description of the notion of "planning freedom restricted by environmental considerations"

Freedom is a generally understood notion that is sometimes (without awareness of it) also used in planning practice. However, there is no clear definition of that notion in the spatial planning literature, and it is not commonly applied in planning practice. Planning freedom, determined by the necessity for environment protection, has been defined according to the following formula.

freedom	+	spatial planning	=	planning freedom	+	science of environment protection	=	planning freedom restricted by environmental considerations

Fig. 7.1 Formula used in development of the notion of "planning freedom restricted by environmental considerations"

Based on an analysis of notions according to the formula (Fig. 7.1) planning freedom restricted by environmental consideration was defined as freedom of planning activities determined by the existence or nonexistence of limitations in land disposal resulting from the presence of different forms of environment protection (Cymerman and Kurowska 2002). The greater the freedom, the fewer the limitations, and thus we have more options in selecting the function and means of development of the areas.

4 Determination of the influence on planning freedom of limitations resulting from protection

Determination of the frameworks for planning freedom restricted by environmental considerations aims to identify the different power that limit the future development of a given area. Determination of the measure of planning freedom restricted by environmental considerations was accomplished as described below.

4.1 Taking inventory of the forms of environmental protection with determination of the scope of protective limitations – the type and scope of limitations were determined based on analysis of legal regulations on protection.

4.1.1 Type of limitations

The limitations are of different natures and scopes, applying to space that is strictly determined and defined, or space that has no specified borders. Considering the above conditions, it is proposed to divide the protection limitations into:

- Protection limitations concerning space without indication of its borders defined as "limitations without spatial location".
- Protection limitations concerning space with specifically indicated borders defined as "limitations with spatial location".

4.1.2 Scope of limitations

The principles of the economic use of real properties on which legal forms of protection are established can be basically divided into bans and orders. Within these frameworks it is also possible to identify limitations that refer to specific forms of space use only. Limitations are divided into four economic-spatial categories: limitations concerning the choice of the function, terrain method of development, principles of use of the real property and activities carried out on the real property.

4.2 Determination of the magnitude of influence of the individual protection forms on planning freedom – questionnaire-based survey

A survey, using an experimental methods – the questionnaire-based survey – was conducted to determine the influence forms of environmental protection on the planning freedom restricted by environmental considerations. The survey was conducted within the framework of a research project (Kurowska 2002).

Table 7.1 List of power levels in environmental protection (Act Environment Protection Law 2001)

Score points	Form of protection
10	National park
9	Nature sanctuaries, Natura 2000 areas, water sources and intake locations
8	Landscape parks, natural monuments, botanical and zoological gardens, protective zones of water intake locations
7	Locations of plants and animals, areas of extraction of mineral resources
6	Protected landscape areas, documented locations, organic soils, forests, cultural sites, spas, protective forests
5	Eco-corridors, protective areas (around nature parks, sanctuaries, landscape parks), Land under organic use, natural-landscape complexes, mineral soils, areas of limited use
4	Communal (rural) parks

5 Development of the technique for making the maps

Planning studies are developed on base maps of different scales. The spatial nature of planning activities and the possibility of their cartographic presentation means that maps are the basic planning material and the best format for presentation of phenomena.

Development of the process for production of maps of planning freedom restricted by environmental considerations was accomplished in three consecutive stages:

- Development of the method for evaluation of space from the perspective of planning freedom restricted by environmental conditions;
- Determination of technical parameters of the map of planning freedom restricted by environmental considerations;

- Determination of technical aspects of the proposed map.

5.1 Proposal of the method for evaluation of space

The proposed method for identification of zones with different planning freedom, classes with different powers of protective restrictions, is a process that is implemented in two subsequent stages: taking inventory and valorisation.

Stage I – taking inventory of environmental restrictions involves an initial review of the studied area (gathering information on the existing forms of protection within the given area and gathering existing documentation). The aim of taking the inventory is to identify the areas covered by protection where limitations in management may exist, influencing the choice of the area function. The inventory can be taken by means of a direct or indirect method.

Stage II – the valuation of protective restrictions is, to a significant extent, the equivalent of area qualification, the determination of the suitability of an area for a specific function – in this case the identification of areas representing different levels of planning freedom. The evaluation of a given area is performed in basic fields. The area valuation can be conducted by applying two criteria for dividing the area into basic fields:

- It is assumed that the smallest unit of territorial division of the country, the registry lot, is the basic field.
- It is assumed that the basic field is the square (e.g. with a side length set at 4 cm) as a geometric figure resulting from artificial division into smaller parts.

Evaluation of the level of planning freedom restricted by environmental considerations requires allocation of the scored points from Table 7.1 to the identified evaluation fields. Individual forms of protection may occupy part of the evaluation field but it is also possible that within one basic evaluation field different protection forms can overlap (partly or entirely). Given the above, it is proposed to introduce the share of the space taken by a given protection form as the weight. The total planning freedom restricted by environmental considerations within the given evaluation field will be seen as the sum of scores for the individual forms of protection present within the evaluated field, taking into account the share of the given form of protection within the entire evaluation field. It may be that the score of the field is higher than 10 points, thus, it is difficult to define an upper limit for Class I.

The choice of basic fields for the valuation of protective restrictions depends mainly on the purpose the map is to serve and the available base maps. Planning freedom restricted by environmental considerations for individual basic fields is computed according to the formula (1):

$$Sp = \sum_{i=1}^{n} (Wo_i * Wp_i) \quad (1)$$

where:

Sp – magnitude of planning freedom restricted by environmental considerations in the studied field

n – number of different forms of protection in the field studied

Wo_i – points scored for protective limitations (within the range of 1 to 10) – (Table 7.1)

Wp_i – share of the area occupied by the given form of protection in the area of the given basic field (within the range of 0 to 1 – depending on the share percentage)

The MapInfo Professional application identifies the size of the object within the basic field and attributes the appropriate value within the 0 to 1 range (where 0 – protected object not present, 1 – protected area equals the basic field area – 100%).

After scoring, individual primary fields are assigned to one of the identified classes of planning freedom according to the principles presented in the following Table 7.2.

Table 7.2 Classes of planning freedom restricted by environmental considerations

Score	Freedom class	Description of planning freedom level	Power of restrictions
≥ 8,1	Class I	no planning freedom	very large protective restrictions
6,1 – 8,0	Class II	very low planning freedom	large restrictions
4,1 – 6,0	Class III	low planning freedom	moderate restrictions
2,1 – 4,0	Class IV	moderate planning freedom	low restrictions
≤ 2,0	Class V	great planning freedom	very few restrictions

5.2 Determination of the technical parameters

The map of planning freedom restricted by environmental constraints is to be a thematic map used for spatial planning purposes. Technical parameters should correspond to the requirements of planning studies and elaboration of thematic maps. This applies mainly to the map scale, base map contents and the form of presentation.

5.2.1 Map scale

The map should be elaborated for the basic level administrative unit area in the base map system, or as a single sheet of the unit (area) map covering that administrative unit.

The basic criteria for choice of map scale are the intended use and role of the map, level of detail and legibility. To allow the use of the elaborated map as baseline material for planning studies, it must be compatible with those studies, and the map contents should present the location of individual phenomena and their intensity (Table 7.3). Thus, the map scale depends on:

- type of planning study;
- level of intensity and type of environment protection forms present;
- size of the area covered by the study.

Table 7.3 Proposed scales of the map of planning freedom restricted by environmental conditions (for a commune)

Type of planning study	Scale of maps for planning studies	Examples of the scale of the planning freedom map
Study of conditions and direction of spatial development of the commune (shaping the spatial policy)	1 : 25 000 1 : 10 000	1 : 25 000 1 : 10 000
Local spatial development plan (local laws)	1 : 5 000 1 : 2 000	1 : 5 000 1 : 2 000
Decisions about conditions of land development for construction and other development	1 : 1 000 1 : 500	1 : 1 000

5.2.2 Base map contents

Planning (location) decisions are taken at different levels and in different situations. Thus, the planning freedom restricted by environmental considerations should be considered in a different way at the level of national studies, voivodeship studies and at the level of communal studies. This diversification results from the scale of problems and the area in which they are present. In this chapter, the commune is the subject of considerations. The presented proposal for determination of areas with different planning freedom levels will apply to the area of a commune with the possibility of using the results, first of all, in planning studies concerning the commune.

The principal map (in classic or digital format), derives from the principal map (thematic overlays) or the topographic map that is used for the development of the proposed map at the scale of 1:5000. The use of topographic maps as the base maps for elaboration of the map of planning freedom restricted by environmental constraints, is proposed, in the scale of 1:10000 and 1:25000.

5.2.3 Form of presentation of the phenomena forming the contents of the map

The phenomenon of planning freedom restricted by environmental considerations is of spatial (surface related) character and thus surface marks delineated in their hachuring or colour are proposed for presentation. The process of developing the map will depend on the base map available (in digital or classic format) (Mazur 2013).

5.3 The map elaboration process

Elaboration of the map should proceed according to the following pattern:

- Making the choice of base map with appropriate scale and contents.
- Choice of the basic field for evaluation and division of the map into the basic fields.
- Evaluation of magnitude of planning freedom restricted by environmental constraints in identified fields and allocation of the appropriate classes.
- Representing the identified classes on the base map.

The map may be produced in two formats, as a traditional analogue map, or in digital format using modern information technologies.

GIS allows, first of all, performance of the assumed spatial analyses and development of different models of phenomena and simulations of processes taking place in the environment. The baseline data for GIS usually consists of source information in the form of different thematic maps, remote sensing images, the results of direct land survey and other types of information in numeric and text formats (Crampton 2009). The results of the analyses, simulations, models or enquiries to the database can be presented in the format of a map, graph or table. GIS is characterised by openness, mobility and the potential for continual extension and updating, which surely facilitates the planning process at local, regional, national as well as global level. An active approach and software tools such as decision support systems (DSS) will facilitate the decision-making process in spatial planning (Menzel et al. 2012). MapInfo Professional possesses functions allowing connection with WMS (Web Map Services) and WFS (Web Feature Services). Thanks to those functions of the programme it is possible, for example, to use the resources of the Polish Geoportal and show protected areas against that background. The notion of planning freedom restricted by environmental constraints has been defined in a much broader way, however, and it covers all naturally valuable areas, as well as areas of restricted use. Thus, the technique used in development of the maps of planning freedom restricted by environmental constraints requires complementing and expanding the scope of information by providing missing source data.

6 Presentation of the potential for using the map in planning studies (determination of usability and application)

The map of planning freedom restricted by environmental constraints illustrates the strength of limiting influence resulting from the presence of legal forms of environment protection in a given area on the choice of the area function. This is a map that belongs amongst maps of the natural environment and the group of environment protection maps.

The map of planning freedom restricted by environmental constraints is intended for use in spatial planning processes, in particular as baseline material for the elaboration of planning studies. As a thematic map, it should represent basic source material for the development of any planning studies (Table 7.4). The possibilities and scope of its use were identified based on the analysis of contents of the individual planning studies. The analysis covered planning studies at the level of the country, voivodeship, county and commune, representing in their contents both the spatial policy and the tools for implementation of that policy.

Table 7.4 Recommendations for using the map of planning freedom restricted by environmental conditions at the commune level

Type of planning study	Possibilities of using the map of planning freedom restricted by environmental conditions
Study of conditions and direction of spatial development of the commune (shaping the spatial policy) Local spatial development plan (local laws)	- in selection of the intended area use (function), - in selection of locations for protected areas, - in selection of locations for industrial and service areas, - in selection of areas for development by construction for organised investment activities, - in selection of locations for development by housing construction, - in selecting the directions of development of transport and technical infrastructure, - in selecting areas for implementation of the tasks resulting from the spatial policy of voivodeships and the state, - in selecting land for afforestation, - in determining the principles for division into land lots, - in determining the line of development by construction, - in designing borders
Decision on conditions of land development by construction and land development	- in determining the conditions for land development by construction and land development

7 Development of the map of planning freedom restricted by environmental constraints for a selected object of research

The technique developed was applied to a selected rural area. In selecting the object for the research project different conditions were considered (Geoffrey 1968). The diversity of forms of environmental protection present, diversification of suitability of soils for production and the rich culture of the region were the major criteria considered. These criteria were met by the commune of Łyse in the County of Ostrołęka in Mazowieckie Voivodship.

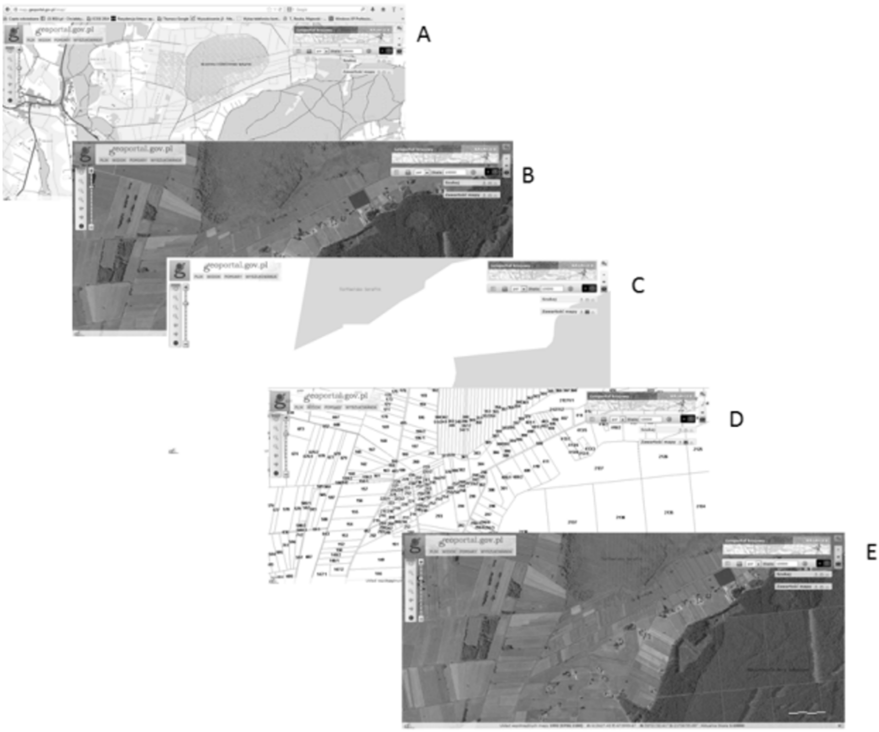


Fig. 7.2 Application content of thematic maps from national Geoportal in Poland
A – topographic map, B – orthophoto map, C – protected area map, D – cadastral map, E - the map of planning freedom

The Polish online resources don't include all the data needed to draw up a map of planning freedom. Figure 7.2 shows examples of thematic maps obtained from national Geoportal. The technology for drawing maps of planning freedom restricted by environmental constraints requires complementing and expanding the scope of information by providing the missing source data.

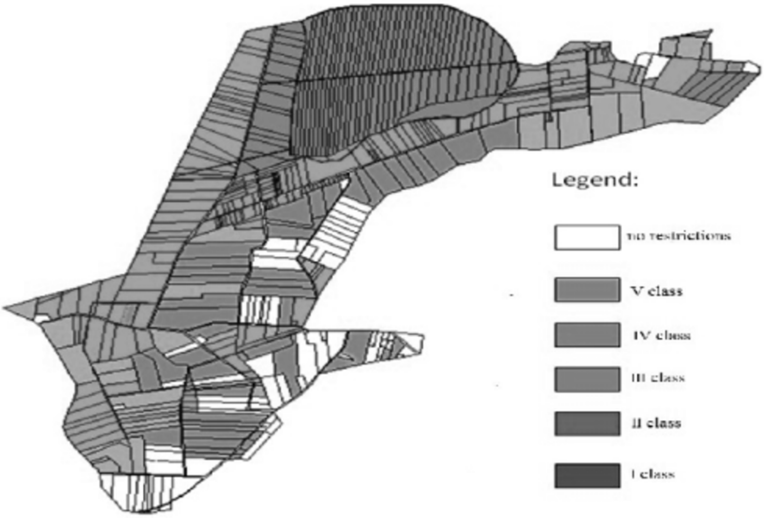


Fig. 7.3 Drawing of the map of planning freedom restricted by environmental constraints (Serafin area, the County of Ostrołęka, Mazowieckie Voivodship)

The map of planning freedom developed in MapInfo Professional for addition of the missing information in relation to protected areas (soil classification, soil type, water intakes, etc.). Figure 7.3 is an image of the planning freedom restricted by environmental considerations map. For the purpose of elaboration of the planning freedom map for the Serafin area the registry lot was used as the basic field. Other thematic layers were not presented so as to maintain the legibility and transparency of the image.

8 Final conclusions

Planning freedom restricted by environmental constraints is a new notion that has not yet been applied in the planning and environment protection practice. The introduction of planning freedom restricted by environmental constraints into those areas would allow an objective perception of the environment protection issue and retention of the most valuable creations of nature. Spatial planning is one of the most effective tools in natural environment protection, and the local spatial development plan is the study supporting that process.

Considering the above observations, the results of studies, analysis of the literature, and our own observations, the following conclusions were drawn:

- Progressing socioeconomic development and the current status of the natural environment require the search for new instruments securing sustainable development. The introduction of the new notion of "planning freedom restricted by environmental constraints" into spatial planning and environment protection would allow objective perception of the issue of rational management of natural environment resources at both local and macro-space levels.

- Various legal regulations concerning environment protection introduce numerous restrictions that are diverse in their scope and coverage. Separate consideration requires not only prolonged time but does not offer any guarantee of their correct interpretation. Development of the method of "summing up different restrictions" and presenting them in the form of classes of planning freedom restricted by environmental constraints will allow objective consideration of environment protection in the processes of spatial planning and facilitate the location of areas offering different levels of planning freedom restricted by environmental constraints.
- The proposed method of determining the areas of different planning freedom restricted by environmental constraints is of a universal nature and allows application at all levels of spatial planning, using modern GIS technologies.
- The current techniques of creating digital maps (existing software) are also applicable to creating thematic maps, in this case the map of planning freedom restricted by environmental constraints. These maps may not only be a tool for gathering and visualising information but they may also facilitate positioning specific functions within space.
- The process developed for the elaboration of the maps of planning freedom restricted by environmental constraints was verified through use of a selected research object. The results are positive opinion as concerns its suitability for application in planning studies.
- The elaborated (sample) map of planning freedom restricted by environmental considerations as a thematic map should be classified in the complex of maps of the natural environment, the group of environment protection maps.
- GIS and DSS (Decision Support System) technologies allow spatial analyses to be conducted considering multiple criteria, leading to identification of areas searched for, in this case areas allowing high levels of planning freedom.

References

Act Environment Protection Law (2001) No 62, item 627

Alterman R, Page JE (1973) The ubiquity of values and planning process, Plan Canada 13(1)

Ažman I, Petek T (2009) Spatial data infrastructure at the surveying and mapping. Proceedings of GSDI 11 World Conference Spatial Data Infrastructure Convergence: Building SDI Bridges to address Global Changes, Rotterdam, 15-19 June 2009

Cartwright TI (1973) Problems, solutions and strategies: a contribution to the theory and practice of planning, Journal of the American Institute of Planners 39(3)

Cymerman R, Kurowska K (2001) Environment protection limits in physical planning. Proceedings of Rural Management and Cadastre Conference, Warsaw, 22-24 June 2001, pp. 201-208

Cymerman R, Kurowska K (2002) Sozoswoboda planistyczna jako instrument ekorozwoju obszarów wiejskich (Planning freedom as an instrument for sustainable development of rural areas), in: Ekoinżynieria dla ekorozwoju, Inżynieria Ekologiczna No. 6, Warsaw, pp. 96-103

Crampton JW (2009) Cartography: maps 2.0, Progress in Human Geography 33(1), pp. 91-100

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

Dukaczewski D, Bielecka E (2009) Comparative analysis of national Geoportals in Europe, Roczniki Geomatyki, T. VII. Z. 6 (36), pp. 35-58

Gans HJ (1970) From urbanism to policy – planning, Journal of the American Institute of Planners, 36(4)

Geoffrey V (1968) The use of speculation, Journal of the American Institute of Planners, 34(1)

Hopfer S, MacEachren AM (2007) Leveraging the potential of geospatial annotations for collaboration: a communication theory perspective. International Journal of Geographical Information Science 21, 921-934

MacEachren AM, Cai G, Sharma R, Rauschert I, Brewer I, Bolelli L, Shaparenko B, Fuhrmann S, Wang H (2005) Enabling collaborative geoinformation access and decision-making through a natural, multimodal interface. International Journal of Geographical Information Science 19, 293-317

Menzel S, Nordström WE, Buchecker M, Marques A, Saarikoski H, Kangas A (2012) Decision support systems in forest management: requirements from a participatory planning perspective. European Journal of Forest Research131, pp. 1367-1379

Mazur M (2013) Cartographic Method of Research as Source of Information in Rural Studies, in: Acta Universitatis Lodziensis, Folia Geographica Socio-Oeconomica 13, pp. 41-58

Kurowska K (2002) Research project KBN no 8 T12E 022 21. Development of the structure of rural areas depending on the protection conditions. Report, University of Warmia and Mazury in Olsztyn

Peng S, Fu GYZ, Zhao XH, Moore BC (2011) Integration of environmental fluid dynamics code (EFDC) model with geographical information system (GIS) platform and its applications, Journal of Environmental Informatics, 17(2), pp. 75-82

Rannestig E (2009) The Swedish National Geodata Strategy and the Geodata Project, Proceedings of GSDI 11 World Conference Spatial Data Infrastructure Convergence: Building SDI Bridges to Address Global Changes, Rotterdam, 15-19 June 2009

Rodriguez A, Abad P, Alonso JA, Sanchez A, Gonzalez C, Mas S, Diez E, Soteres C, Potti H (2009) Data and Services Availability in Spanish NSDI, GSDI 11 World Conference Spatial Data Infrastructure Convergence: Building SDI Bridges to address Global Changes, Rotterdam, 15-19 June 2009

The Solution of Massive Tasks in GIS Exemplified by Determining Terrain Surface Types above Gas Pipelines in the Czech Republic

Dalibor Bartoněk, Jiří Bureš, Irena Opatřilová[1]

Abstract The automated creation of cartographic products made with the software tools of geographic information systems (GIS) allow people to solve extensive and complicated tasks. These tasks lead to massive calculations which are very demanding in terms of data preparation, machine time and the memory capacity of hardware. This chapter deals with theoretic analysis of the complexity of tasks and the possibility of optimisation of sub-processes which lead to acceptable solutions. The quality of results achieved is also discussed here. Theoretical assumptions were verified through a data analysis project involving the storage of gas facilities under certain types of surface terrain in the Czech Republic. This analysis was performed in order to determine the re-built costs of gas facilities (pipelines) and the valuation of necessary costs in building new networks. The authors undertook this project for the GasNet, Ltd. Company which is a part of a RWE group in the Czech Republic. The results have general significance for the creation of cartographic products with GIS support.

1 Introduction

In GIS tasks the vast majority of operations are performed with the geo-database which contains territorial located data (geo-data). GIS processes are distinctly different according to their range of modelled area and the type of the input data from the other processes of the common information systems. High demands are made of output quality during map production using GIS technology. The large amount of input data and an extensive range of GIS tools are necessary to achieve

D. Bartoněk, J. Bureš, I. Opatřilová
Institute of Geodesy, Faculty of Civil Engineering, Brno University of Technology, Czech Republic
e-mail: bartonek.d@fce.vutbr.cz, bures.j@fce.vutbr.cz, opatrilova.i@fce.vutbr.cz

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_8,

the required quality. This leads to complexity in the whole process. The time demands of these types of tasks is analysed in this chapter, according to the type of input data and a scale factor. The scale factor is a ratio of the extent of the modelled area to the dimension of the smallest detail in the input data. Problems with complicated tasks in GIS were verified in a project of data analysis for the storage of gas facilities under certain types of surface terrain in the Czech Republic, covering an area of 64,350 km^2. This analysis was performed in order to determine the re-built costs of the gas facilities (pipelines) of the RWE energy group and the valuation of costs which would be necessary for building new networks. The results are used for strategic planning of maintenance and the development of gas facilities in the Czech Republic.

2 Related works

The issue of complicated tasks has been researched since the 1970s, and there are many publications on this theme. We consider only those works that are directly related to our chapter. Many methods were developed for processing massive tasks, and they can be divided into these categories:

1. Use of parallel processes in the task which can be processed simultaneously. This group is characterised by methods requiring special hardware, mainly consisting of multiprocessor computers. These computers are also called accelerators. We can include the following papers in this category: Juany et al. (2013), Weiss and Bailey (2011). Forslund et al. (1992), Faro et al. (2011), Frache et al. (2013), Xie et al. (2008), Bartoněk and Opatřilová (2014).

2. Dividing tasks into sub-parts, the solution of which is simpler and more reliable than solution of the whole task. These sub-parts are solved serially on computers of high performance. We can include the following papers in this category: Wang et al. (2013), Wen et al. (2013), Zhang et al. (2011), Min et al. (2005), Steinfadt (2013).

3. Use of special algorithms which optimise processes in given tasks. Use of discreet simulation belongs into this group. Special algorithms are described in papers, for example: Yang et al. (2013), Sakellari and Loukas (2013). Richter and Döllner (2013), Park et al. (2004), Bartoněk and Bureš (2012).

4. Implementation of organisational measures into computing processes with the aim to simplify and accelerate the time for solution of the given task. In most cases, these methods are based on graph theory. The solution of massive computing using special methods is the content of papers e.g. Chetyrkin et al. (2007), Cheby et al. (2012), Juany (2011), Luo et al. (2005), Yang et al. (2011).

The authors were inspired by works in the second and fourth categories. The reduction of time for solving tasks, memory requirements and thus cost savings, are achieved through a combination of organisational measures and the appropriate distribution of tasks on the principle of territorial administrative units.

3 Theoretical foundations

The solving of tasks in GIS can be described mathematically by this formula:

$$R = (I, S, Q, Y, \delta, \varphi, \psi), \tag{1}$$

where

$$I \subseteq (A_1 \times A_2 \times \cdots \times A_l), \tag{2}$$

I is the set of input datasets, A_i are partial datasets
l is the number of sub-datasets
S is the set of internal states of GIS
Q is the set of acceptable results (final layers) which meet the required quality
Y is the set of outputs from GIS
δ is the coefficient of territorial details given by:

$$\delta = \frac{P_{tot}}{P_{det}}, \tag{3}$$

where
P_{tot} is the total area of the modelled territory, which is solved in the context of the given task,
P_{det} is the area of detail, which is described in input layers I,
φ is the transform (input predicate):

$$\varphi : I \to S \tag{4}$$

This transform is characterized by the preparation of input

$$\psi : S \cap Q \to Y \tag{5}$$

where ψ is the transform (output predicate).
The operation γ: $S \cap Q$ represents the process of output quality (cartographic product), which can be expressed as:

$$\gamma : S \cap Q = f(C, P, T, L, A) \tag{6}$$

where (according to ISO 19 113):
C is the completeness of the input data
P is the positional accuracy
T is the time accuracy (timeliness)
L is the logical (topological) accuracy
A is the attribute (thematic) accuracy.
The time demand of the given task in GIS can be expressed by this formula:

$$T = f(\delta, \varphi, \psi, l) \tag{7}$$

In this equation, δ is a scale of transformation from the reality into the digital geo-database, φ represents the complexity of the preparation of input data, i.e. the transformation of the input layers to the inner states of GIS, ψ has the character of internal processes in GIS, i.e. the character of tools for spatial operations, analysis, editing, checking quality γ etc., and l is the number of types of input layers - see Equation (2). Let us now analyse the time demand of tasks in the GIS according to Equation (7). The parameters in this equation can be divided by the rate of user influence into two groups:

1. The input φ and output ψ predicate are given by the character of the task. The user may affect only a small part.
2. δ and l coefficients, on the contrary, can be affected quite significantly. The parameter n is determined by the success of output and its quality, the coefficient δ depends not only on the processing time, but also on the reliability of internal processes ψ.

Equation (7) can also be rewritten in the form:

$$T = T_{Ia} + T_{Im} + T_{Oa} + T_{Om} \tag{8}$$

where
T_{Ia} is the time for the automatized preparation of data
T_{Im} is the time for data preparation, which takes place semi-automatically or manually (the input control)
T_{Oa} is the time for data processing and automatic calculation of results T_{Om} is the time for data processing and calculation of results manually or semi-automatically (the output control).

The authors have empirically determined the value of the function in Equations (7) and (8) through experiments within the pipeline project.

4 Experimental results

The GIS time demands of the task were verified within a data analysis project involving the storage of gas facilities under certain types of surface terrain in the Czech Republic, covering an area of 64,350 km^2. The aim was to classify the surface above the linear route of the gas line with a maximum error rate of up to 5%. Data was chosen from the Czech Republic due to the extent of its territorial coverage. An orthophoto with a resolution of 25cm/1pixel was selected as the basic dataset for the project. The main method used in the project was supervised classification of raster image using the Maximum Likelihood method in ArcGIS 10.0. It is widely accepted that this classification has an average success rate of 85% (Sijbers et al., 1998), which is insufficient in our case. The input data was therefore supplemented by a dataset of nine layers of communications, which is a part of the fundamental basis of geographic data of the Czech Republic (ZABAGED). The preliminary investigation showed that the largest error in raster image classification is where vegetation obscures the surface of roads and distorts the classification result. The quality of the automated classification was therefore refined by approximately 10% following reclassification using the vector dataset of nine layers of the communications from ZABAGED database as a filter.

The automated technological line in ArcGIS was developed for processing classification. This line includes:
1. The preparation of sub-datasets by selection from a data warehouse into a template of the ArcGIS project. This project covers the territorial unit of the municipality with extended state administration (ORP), of which the average area is 368 km^2.
2. Spatial analysis with a dominant process of supervised classification via the Maximum Likelihood method and filtering of results according to layers of vector datasets of communications from ZABAGED.
3. Export of processed results for a territorial segment of ORP into text and a numeric contingency table in Excel, and into SHP graphical format with the classified vector route of the pipeline.

Control scripts were created in Python 2.6 with support of ESRI libraries for geoprocessing for all 1-3 processes. These operations are performed fully automatically with the exception of manual visual inspection with subsequent editing of the results between the sub-processes 1 and 2, and 2 and 3.

Table 8.1 presents some of the results of the spatial analysis of gas facilities stored under certain types of surface terrain in the Karlovy Vary region over an of 3,315 km^2, where an average extent of the ORP sub-processing units is 474 km^2. Table 1 introduces a view of the percentage portions of the length of gas lines stored under the surface terrain of Type 1 (paved surface) and Type 2 (unpaved). If these percentage portions are converted by the relevant economic factors, such as unit cost of the reconstruction of the type of pipeline route (low-pressure + medium-pressure, service pipe, high-pressure) under the appropriate type of surface, it is possible to deduce the total investment costs of reconstructing the pipeline route in the area.

Table 8.1 The percentage of storage of the gas line route under types of the surface terrains

Area of ORP	**Low-pressure and medium-pressure pipe [%]**		**Service pipe [%]**		**High-pressure pipe [%]**	
Average area of 474 km^2	Type 1	Type 2	Type 1	Type 2	Type 1	Type 2
Aš	51,1	48,9	39,7	60,3	1,0	99,0
Cheb	45,9	54,1	33,6	66,4	0,9	99,1
Mariánské Lázně	47,5	52,5	37,2	62,8	1,3	98,7
Karlovy Vary	60,4	39,6	45,9	54,1	1,1	98,9
Ostrov	51,2	48,8	36,5	63,5	0,8	99,2
Kraslice	67,0	33,0	44,4	55,6	0,6	99,4
Sokolov	54,3	45,7	42,9	57,1	1,0	99,0
Total for area of region (area of 3,315 km^2)						
Karlovy Vary region	54,2	45,8	40,7	59,3	1,0	99,0

Note: Type 1 – paved surface, Type 2 – unpaved surface

The visualised example of the ORP territory data project (left), including schematic view of a data structure and detail of a processed orthophoto image along a route of the gas line in the sub-categories, is in Figure 8.1.

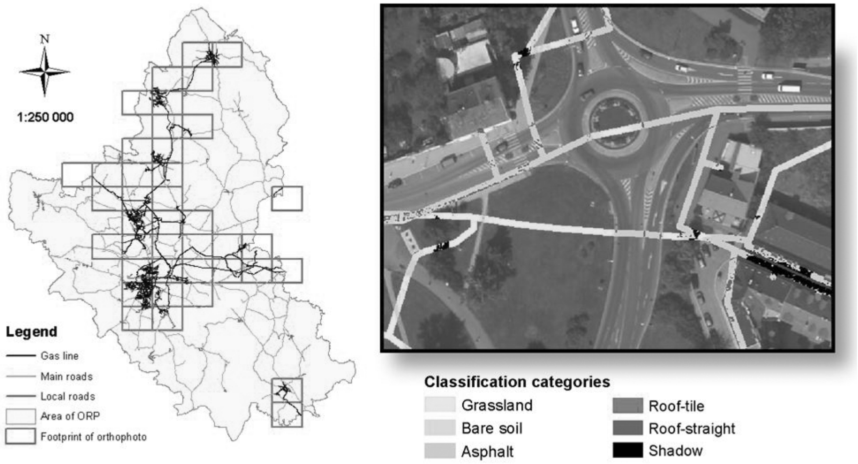


Fig. 8.1 Example of an ORP data project and a detail of image classification along route of the gas line

The example of zoom to detail in the final processed image of route of gas line is in Figure 8.2. Figure 8.3 shows percentage portions of automated sub-processes:

data preparation (T_{Ia}), data processing (T_{oa}) and the process of manual work (T_{Im}+T_{Om}) for the entire studied territory with a total area of 64,350km^2. Data volume was 388 GB, while the average processing unit for an area of the ORP was 368 km^2 (about 2.1GB of data). The automated processing took place on a PC with Intel Core i5 CPU, 16 GB DDR 1600 MHz, HDD SSD with NVIDIA GTX650 2 GB graphic adapter and using Microsoft Windows 7 64-bit version. The processing software was ArcGIS 10.0 and Python 2.6, both 32-bit versions.

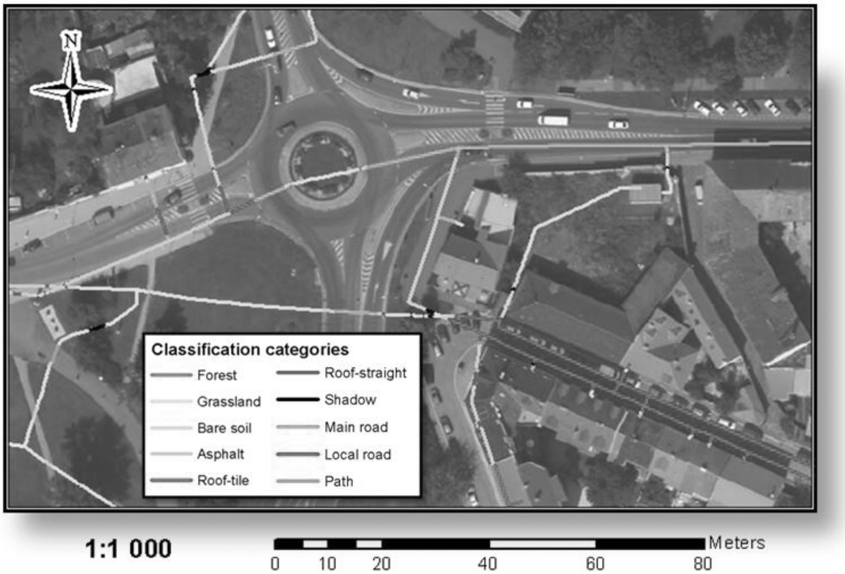


Fig. 8.2 The detail of final processed image of route of gas line

Although the manual work takes proportionally more than half the time, high automation of data preparation and data processing in the form of Python scripts greatly streamlines the entire process of evaluation.

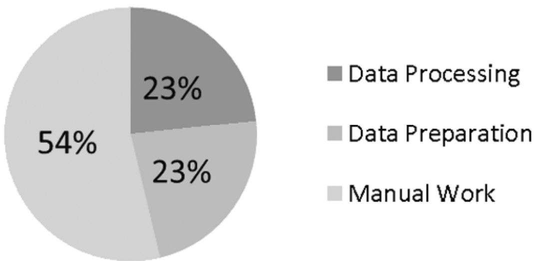


Fig. 8.3 Portions of automated sub-processes and manual work for the whole task

The time demand percentages for manual processing only, data processing through Model Builder and processing through a Python script are illustrated in

Figure 8.4. Manual processing is not automatised and is time-consuming; ModelBuilder makes similar time demands. Processing through ModelBuilder is semi-automatised. Data processing using Python script is the fastest, and most effective and is fully automatised.

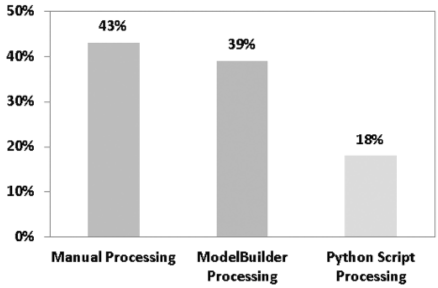


Fig. 8.4 Portions of time demand for different types of data processing

5 Conclusions

The aim of this chapter was to present an analysis of the complexity of the tasks and possibilities of optimising the sub-processes and contexts that lead to an acceptable solution of massive tasks in GIS, and the results of a data analysis using geodata from the Czech Republic covering an area of 64,350 km^2. We analysed common processing of orthophoto datasets (raster, 25cm/1pixel), ZABAGED (vector, 9 levels) and gas lines (vector, 3 levels). The volume of data was 388 GB total. The results of the processed analysis confirmed that:

1. The effectiveness of data analysis for a large territory depends on the rate (grade) of automation. The time demand ratio achieved was really on 23% for preparing data (fully automated), 23% for data processing (fully automated) and 54% for manual work. Manual work is not completely ruled out in any process. Fully automated data processing is more than twice as fast than the equivalent fully manual solution.

2. Technical reliability of processing depends on the coefficient of territorial details δ (has the character of scale) expressed by Equation (2). In our case, the distribution on the data units for processing with the size of territorial unit of the ORP (the municipality with extended state administration) was showed as suitable for total processed territory with the area of 64,350 km^2. The ORP has the average area of 368 km^2 and there is $\delta \approx 6.10^9$. The use of smaller territorial unit leads to an excessive increase of cycles of sub-processes and the use a larger territorial unit greatly increases the time and spatial complexity of operations and significantly reduces the technical reliability of processing with regard to the limits of computer technology. Tests have discovered that some spatial analyses (raster-vector conversion, etc.) of large areas with a given resolution are beyond the capabilities of current hardware and software (ArcGIS 10.0) (Bureš et al. 2013).

3. Absolute factual reliability (error rate) of data analysis, which is determined through statistical analysis of a representative selective data sample of 20% of the total data volume, does not exceed 5%.

The set of tools that were created for solving the project device, which is generally suitable for tasks of the same class. This means that in the future it should be possible after a certain time to update the classification of surface terrains above gas facilities using the procedure described above.

Acknowledgments The work was undertaken within project FAST-S-13-2069, “Data mining from available geospatial standard data sources using GIS” and within research project MŠMT (Ministry of Education of the Czech Republic) AdMaS ED2.1.00/03.0097 Nr. HS12357021212200 “Data analysis of surfaces above RWE gas pipelines in the Czech Republic”.

References

Bartoněk D, Bureš J (2012) Expert GIS for locating of archaeological sites, 4th International Conference on Cartography & GIS Albena, Bulgaria, June 2012, ISSN 1314-0604

Bartoněk D, Opatřilová I (2014) Design of man–machine interface for mobile mapping. Advanced Science Letters, Volume 20, Number 2, pp. 501-504 (4)

Bureš J, Bartoněk D, Opatřilová I (2013) Data analysis of surfaces above RWE gas pipelines in the Czech Republic. Project AdMaS ED2.1.00/03.0097 Nr. HS12357021212200. Brno University of Technology, Czech Republic (in Czech)

Cheby J, Wang W, Gao R (2012) Massive RDF data complicated query optimization based on MapReduce. Physics Procedia, Volume 25, pp. 1414-1419

Chetyrkin K - G, Kühn J - K, Sturm C (2007) Recent progress in computing four-loop massive correlators. Nuclear Physics B - Proceedings Supplements, Volume 164, pp. 203-206

Faro A, Giordano D, Majorána F (2011) Mining massive datasets by an unsupervised parallel clustering on a GRID: Novel algorithms and case study. Future Generation Computer Systems, Volume 27, Issue 6, pp. 711-724

Forslund D - W, Hansen C, Junker P, John W - S, Tenbrink S, Breton J. (1992) High-speed networks, visualization, and massive parallelism in the Advanced Computing Laboratory. Computing Systems in Engineering, Volume 3, Issues 1–4, pp. 521-524

Frache S, Chiabrando D, Graziano M, Vacca M, Boarino L, Zamboni M (2013) Enabling design and simulation of massive parallel nanoarchitectures. Journal of Parallel and Distributed Computing. Vol. 74, Issue 6, pp. 2530–2541

Juany B (2011) A Short Note on Data-Intensive Geospatial Computing. in Information Fusion and Geographic Information Systems. Lecture Notes in Geoinformation and Cartography 2011, pp 13-17

Juany L, Tang G, Liu X, Song X., Yang J, Liu K (2013) Parallel contributing area calculation with granularity control on massive grid terrain datasets. Computers & Geosciences, Volume 60, pp. 70-80

Luo Y – W, Wang X – L, Xu Z - Q (2005) An agent approach to spatial information grid architecture design. Computing and Informatics, Volume: 24, Issue: 2, pp. 201-222

Min J, -K; Park H, - H, Chung C - W (2005) Multi-way spatial join selectivity for the ring join graph: Information and Software Technology, Volume: 47, Issue: 12, pp. 785-795, DOI: 10.1016/j.infsof.2005.01.002

Park H - H, Min J - K, Chung, C - W, Chány T - G (2004) Multi-way R-tree joins using indirect predicates. Information and Software Technology, Volume 46, Issue 11, pp 739-751

Richter R, Döllner J (2013) Concepts and techniques for integration, analysis and visualization of massive 3D point clouds. Computers, Environment and Urban Systems, Volume 45, pp. 114–124

Sakellari G, Loukas G (2013) A survey of mathematical models, simulation approaches and testbeds used for research in cloud computing. Simulation Modelling Practice and Theory, Volume 39, pp. 92-103

Steinfadt S (2013) Fine-grained parallel implementations for SWAMP+ Smith–Waterman alignment. Parallel Computing, Volume 39, Issue 12, pp. 819-833

Sijbers J, den Dekker A – J, Scheunders P, Van Dyck D (1998) Maximum likelihood estimation of Rician distribution parameters. IEEE Transactions on Medical Imaging 17 (3): 357-361. doi:10.1109/42.712125. PMID 9735899

Xie Z, Wu L, Ye Z (2008) A Prototype Design of Parallelizing GIS operations. Conference: Geoinformatics 2008 and Joint Conference on GIS and Built Environment - Advanced Spatial Data Models and Analyses Location: Guangzhou, Republic of China, Proceedings of the SPIE, Volume 7146, pp. (2009), article id. 71462F, 9 pp

Yang C, Goodchild M, Qunying H (2011) Spatial cloud computing: how can the geospatial sciences use and help shape cloud computing? International Journal of Digital Earth. Volume: 4 Issue: 4, pp. 305-329, Article Number PII 938806602 DOI: 10.1080/17538947.2011.587547

Yang D, Feng Y, Juan Y, Han X, Wang J, Li J (2013) Ad-hoc aggregate query processing algorithms based on bit-store for query intensive applications in cloud computing. Future Generation Computer Systems, Volume 29, Issue 7, pp. 1725-1735

Wang L, Tao J, Ranjan R, Marten H, Streit A., Chen J, Chen D (2013) G-Hadoop: MapReduce across distributed data centers for data-intensive computing. Future Generation Computer Systems, Volume 29, Issue 3, pp. 739-750

Weiss B, Bailey M (2011) Massive Parallel Computing to Accelerate Genome-Matching. GPU Computing, Gems NVIDIA Press, pp. 173-184, DOI: 10.1016/B978-0-12-384988-5.00012-7

Wen, J., Ma, Y., Liu, P., Sun, S. (2013). Distributed multipliers in MWM for analyzing job arrival processes in massive HPC workload datasets. Future Generation Computer Systems, ELSEVIER, http://dx.doi.org/10.1016/j.future.2013.12.009

Zhang W, Zhao M, Tu Z, Lou S (2011) Research and Implement of Distributed Nodes Collaboration-Based Management and Publishing Technologies for Massive Geospatial Information. Massive Computing. Advances in Intelligent and Soft Computing 2011 Vol. 112, pp. 391-400

Approaches to Infrastructure Provision of Spatial Water-ecological Data on the Ob` River Basin Systems

Irina N. Rotanova and Olga V. Lovtskaya[1]

Abstract This chapter considers issues in the infrastructure provision of spatial water and ecological data about the river basin systems. The creation of a unified information space implies solving some organisational, specific and infrastructure-related tasks. The integration of disembodied databases, through the development of a common infrastructure for spatial data, is among the objectives. A scheme for the unified information space should represent its organisational levels: basic, departmental and subject ones. Much attention is given to basic spatial data, but thematic spatial data (specific and interdisciplinary) also deserves consideration. The creation of a common infrastructure for spatial water and ecological data for a large river basin system is discussed. The principles of spatial data infrastructure formation and the informational-cartographic support of geoinformation-analytical system "Water and Ecology of Siberia" as a component of the decision support system are aimed at the steady functioning of Schemes of Complex Use and the Conservation of Water Objects in the Ob' basin. Examples of thematic river basin system maps are presented.

1 Introduction

New possibilities for the integration and analysis of a large body of varied spatial data and its use in decision support systems (DSS) depend on the preconditions available:

- advanced information technologies,

I.N. Rotanova
Altai State University, Institute for Water and Environmental Problems, Siberian Branch of the Russian Academy of Sciences
Barnaul, Russia
e-mail: rotanova07@inbox.ru

O.V.Lovtskaya
Institute for Water and Environmental Problems, Siberian Branch of the Russian Academy of Sciences
Barnaul, Russia
e-mail: lov_olga@inbox.ru

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_9,

- constant development of concepts for creation of information-analytical systems,
- accepted ideology in the creation of spatial data infrastructure.

Basic spatial data is a leading segment (considered as a specific unified form) of information resources.

Domain data is formed as the result of specific mono-, interdisciplinary and departmental studies.

Spatial data infrastructure should be determined by the goals and objectives of information processing.

The creation of an object-oriented infrastructure of spatial data for large river basins is of great importance. The Ob' basin serves as the basis of a spatial data infrastructure of river basin systems (Koshkarev 2009; RFGD 2006).

2 Statement of the Problem

Experience in the development and implementation of national and international spatial data infrastructures suggests three required constituents:

- basic spatial data,
- standards for spatial data representation and exchange,
- metadata bases (Moellering 2006).

Spatial data infrastructure (SDI) involves sets of spatial data about subject specialisation and geoinformation services that provide users' with access to redistributed spatial data resources, data distribution and exchanges for increasing the efficiency of their creation and use.

The basic spatial data as a part of information-cartographical provision should be in line with the organisational-management and industrial-technological structure. Their infrastructure is aimed at the support of management and technological processes.

A large volume of source space-related data, the necessity in distributed access to geoinformation, the demand for analytical search tools call for the formation of infrastructure of spatial interdisciplinary data for management decision making.

These are as follows:

- *inventory* – the study of the environment and its components, the use of natural resources and natural potentiality of the territory;
- evaluation– the assessment of anthropogenic impacts on the environment, the extent of unfavourable processes, the aftereffects of natural or anthropogenic eco-catastrophes, the ecological-geographical evaluation of the territory;
- *dynamic* – the study of changes in the environmental and natural conditions due to economic and natural factors, nature management and the anthropogenic impact on the environment;

- *predictive* – the forecast of environmental/natural conditions change caused by the anthropogenic impact or the development of natural complexes the revealing of trends and dynamics of natural evolution as a result of human economic activity.

3 Approaches and Discussion

Practical application of GIS projects involves different activities involving:

- development of systems for ecological monitoring on the local, regional and basin levels;
- development of territorial cadastres of natural resources;
- assessment of state and use of natural resources including water, land, forest, biological and atmosphere;
- ecological assessment of problems caused by anthropogenic factors (eco-diagnostics);
- ecological expertise about territories, objects and projects;
- development of GIS technologies for sustainable development of territories as well as for solutions to hydrological, water-resource, water-management and water-related ecological problems.

Schemes of Complex Use and the Conservation of Water Objects (SCUCWO) have been developed for the basins of some large rivers of Russia. In line with normative-legal documentation, SCUCWO contains information support activities, including the mapping and development of basin geoinformation systems, i.e. the creation of an infrastructure of spatial data at a basin level (RF Water Code 2008; RF Government Regulation 2006).

The goal of SCUCWO development is the creation of tools for DSS. Item 12 of "Methodical guidance..." states that schemes are based on GIS technologies in compliance with technical and software requirements to support digital map layers (RF Ministry for Natural Resources 2007a). Item 22 refers to mapping and GIS as fundamental activities under the development of programme options for the conservation and restoration of water objects, the steady functioning of water-economic systems, and the achievement of target indices on prevention of negative water effects.

This approach is of great importance for the provision of uniformity and continuity of information, the normative-methodical systems of the Russian Federation in the field of water object conservation, systematisation of materials on the state of water objects, and the structure of water-economic and water protective activities (Pryazhinskaya et al. 2011).

The schemes must include a set of situation, evaluation, operative and predictive maps (in electronic and paper form) constructed at scale 1:1 000 000 – 1:100 000, and supported by an inset map of larger scale, if necessary (Rotanova et al. 2005, Rotanova et al. 2009) (Fig. 9.1).

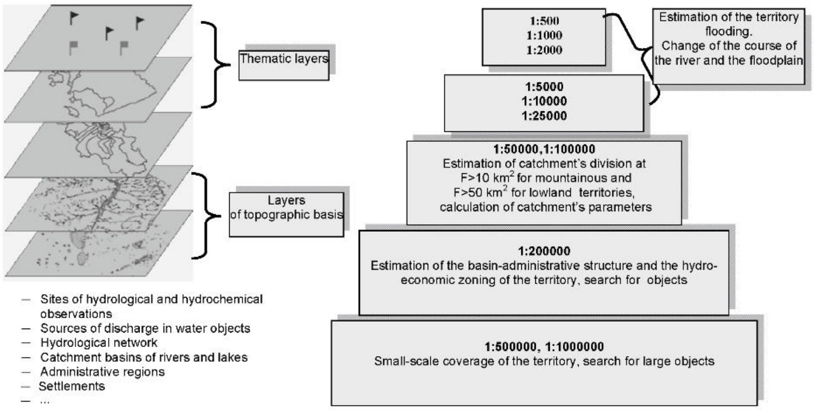


Fig. 9.1 Information Support (geospatial data)

A set of situation maps representing factual information for a moment of their construction consists of:

- a topographic map;
- a landscape map with the mapped protected areas;
- a drainage map with boundaries of hydrographic units and water-resource regions, hydrologic and hydrochemical monitoring stations including tables of studied hydrological situations in the river basin;
- a map of water resource regions including their major characteristics;
- a map of water objects by categories including tables that characterise water objects and their regimes;
- a map of basin infrastructure with water management systems and waterworks facilities including the tables with the parameters and features;
- a map of ground water aquifers;
- a map of aquifers characterised by intensive ground water intake (monitoring wells, deposits of ground water, boundaries of depression whirlpools, aquifers' protection from pollution).

A set of evaluation maps representing the outcomes of the data analysis from situation maps and documentary data on water object management consists of:

- a map of watershed zoning by level of anthropogenic load on water objects;

- a map of water risks stipulated by various water impacts;
- a map of the basin's territories exposed to occasional floods (at different water providing - 1%, 3%, 5%, 10%, 25% and 50%);
- a map of the river basin by level of flood threat;
- a map of key types of water use;
- a map of natural and industrial pollution of surface water;
- a map of natural and industrial pollution of ground water;
- a map of water management balance (by water-resources regions);
- a map of water object assessment based on data from the state hydrochemical monitoring of water objects;
- a map of environmental assessment of water objects;
- a map of protection of exploited aquifers from pollution.

A set of operative and predictive maps that forecast situations consists of:

- a map of predictive changes in water content of the river basin for the period of the scheme validity (taking into account natural-climatic and anthropogenic factors);
- a map of predictive change of anthropogenic load on water objects of the river basin for the period of scheme validity;
- maps of limits and quotas for water intake from water objects according to stages of scheme implementation (by water-resources regions);
- maps of limits and quotas for waste water discharge into water objects according to stages of scheme implementation (by water resource regions);
- maps of target indices of water quality in water objects;
- maps of target indices of negative effects of water;
- maps of development of water objects and systems monitoring;
- maps of planned structural activities to be implemented on the basin's territory;
- a map for forecasting depression whirlpool development within ground water basins and aquifers characterised by intensive ground water exploitation.

Unfortunately, these maps, which differ in quantity and quality, are not methodically supported. Moreover, the indices to be mapped are not approved.

SCUCWO should be developed for one of the largest water objects in Russia- the Ob basin. The watershed area of Ob river (3 million km2) takes 5th place in the world and first in Russia; it is the third (after Yenisei and Lena) in the annual runoff (Vedukhina et al. 2011) (Fig. 9.2).

Fig. 9.2 The Ob' river basin

A complex of research activities precedes SCUCWO development. The goals of research activities are as follows:

- integrated assessment of water objects in the Ob'-Irtysh basin, qualitative and quantitative assessment of surface and ground water;
- elaboration of information-modelling complexes and decision support systems for solving tasks of integrated water resource management in the Ob' basin;
- scientific grounding of methods and tools for steady water use and hydroecological safety;
- information validation for the development of the Ob' basin SCUCWO.

Major tasks in a geoinformation-cartographic block are the following:

- cartographic assessment of the current state of water resources and their use in the Ob' basin:
- collection, processing and analysis of available cartographic source information containing the data on the description and assessment of conditions for water

resource formation as well as on qualitative and quantitative analytical and evaluation indices of water object state in the Ob' basin;

- application of cartographic research methods to the integrated assessment of the state of water objects in the Ob' basin;
- formulation of basic principles and standings of water resource and water ecological cartographic methods for information support under Ob' basin SCUCWO development;
- preparation of basic digital maps and materials with spatial data infrastructure for development of a series of situation, evaluation and predictive maps.

Cartographic investigations of formation processes of surface and ground water's quality and quantity including their influential factors:

- landscape-cartographic field works to obtain data on the environmental assessment of water objects and their catchments;
- method preparation of water-resource and water-ecological small-, mid-, and large-scale mapping for geoinformation-cartographic support under SCUCWO development;
- cartographic evaluation of qualitative and quantitative indices of surface and ground water and their influential factors using the existing cartographic methods;
- structure development of specific databases of the basin's water objects in line with the state water register and monitoring of water objects;
- elaboration of the concept, structure and information content of a geoinformation-cartographic block of DSS for water resource management in the Ob' basin (using model water objects);
- creation of a cartographic block with target GIS database, development of cartographic support for a pilot GIS project.

The creation of scientifically grounded infrastructure of spatial data and information-cartographic support for water management in the Ob' basin due to development of:

- a cartographic method for evaluation of the influence of diffusive sources of water pollution;
- a block of geoinformation-cartographic support for DSS information-modelling complexes aimed at the management of specific water ecosystems in the Ob' basin;
- integrated assessment maps of water objects subject to the peculiarities of watersheds and resource potential;
- a GIS cartographic block for the DSS to manage water resources in the Ob' basin;
- cartographic methods and the construction of water resources and water-ecological maps for model basin rivers.

To study natural conditions and to reveal the peculiarities influencing water resource formation in the Ob' basin, the territory was split into hydrographic units, i.e. river basins (a basin level), sub-basins (a sub-basin level), water resource regions and sites (RF Ministry for Natural Resources 2007b).

According to water- management zoning, the territory of the Ob'–Irtysh basin is divided into 72 water- management section: 36 are in the Ob' basin and 36 in the Irtysh river basin.

Despite uneven coverage of the basin area, a set of landscape-typological maps for specific administrative territories was constructed. At present, numerous approaches to defining the natural territorial complexes of different ranks hamper complex assessment of the region under study.

The scale for construction of original and resulting maps was justified.

The infrastructure of the spatial data for GIS creation was defined. Much consideration was given to the development of basic geodata, the unification of the original information (the creation and use of classifiers), and the development of a custom interface for data retrieval, analysis and visualisation.

The research outcomes served as the basis for the development of a geoinformation-analytical system (GIAS), the "Water and ecology of Siberia". GIAS is considered as the element of DSS information content for water management in the Ob basin (using the model water objects) and the development of territorial systems of Siberia (Lovtskaya et al. 2007).

The specific feature of GIAS is its water-ecological orientation, created catalogue of distributed geoinformation resource metadata and the system developed for further introducing mathematical modelling results and field/reference data

GIAS uses the technologies of remote data processing, the results of cartographical modelling, methods of interdisciplinary data integration and the outcomes of spatial characteristics and features research.

In the framework of GIAS, typical GIS problems and specific tasks are solved:

- the formation of ground waters, physical and chemical characteristics of groundwater flow,
- the analysis of ground and surface water quality, their availability for drinking water supply,
- the characteristics of sources and levels of water pollution,
- the development of rapid/efficient access and information storage in databases, including satellite data ,
- mathematical-cartographical modelling.

The concept of the data warehouse (a subject-integrated, invariant dataset formed to support decision making) serves as the basis for the creation of thematic databases. In terms of multidisciplinary information the formation of the data warehouse assumes an approach that:

- is oriented to ecologically significant objects (water basins, territorial entities) and situations (assessment of state, impact and aftereffects);
- includes the object-oriented datasets containing consistent and aggregated cartographical and factual information.

The objective level of GIAS primarily involves river basins that represent hydrographical units in a specified order hierarchy. Major basins are those of the Ob and Irtysh rivers. In the basins of the principal rivers, large, medium-sized and small river basins are specified. Hydrographic units are used in the analysis of factors for the formation of water resources and the ecological state of water objects (Vedukhina et al. 2009).

The GIAS objects also include RF subjects (the units of political division) and the units of hydrological zoning representing the system of hydrological sites specified with hydrographical-geographical and economic-geographical approaches to the territory zoning (Vinokurov et al. 2012).

The description of attributive and cartographical information and the formation of metadata complies with the national standard on spatial data content (National Standards 2006). Using the standard, its “projection” on the subject domain GIAS “Water and ecology of Siberia” was made, and a subject profile oriented to the system was obtained.

Ecology-oriented maps were integrated into large thematic blocks:

- water-ecological,
- bio-ecological,
- anthropo-ecological,
- socio-ecological,
- economic-ecological,
- integrated.

Within the framework of GIAS "Water and ecology of Siberia" various thematic maps have been created and data has been obtained, in particular, characterising the flow anomalies, as well as natural and industrial pollution, of surface water. The creation of assessment maps of flow anomalies (freezing, drying up) is based on the method of hazard and risk mapping for low water in rivers (Shoygu et al. 2005). These natural phenomena have a significant impact on the possibility of economic use of water bodies and the probability of hydro-ecological risks.

The maps show areas of watersheds with different episodes of the freezing and drying up of rivers. The mapping unit is river catchments. The freezing and drying up of rivers is displayed by the calculation of average monthly water consumption in years with various availabilities of water (shallow, medium and high-water). The hydrographs have been calculated, describing the annual distribution of runoff for hydrological supporting alignments that have been used to assess the area.

The maps are made for basins of water objects, characterised by these processes and regularly observed; for example, for the Ishim River (Fig. 9.3) (Rotanova 2011a).

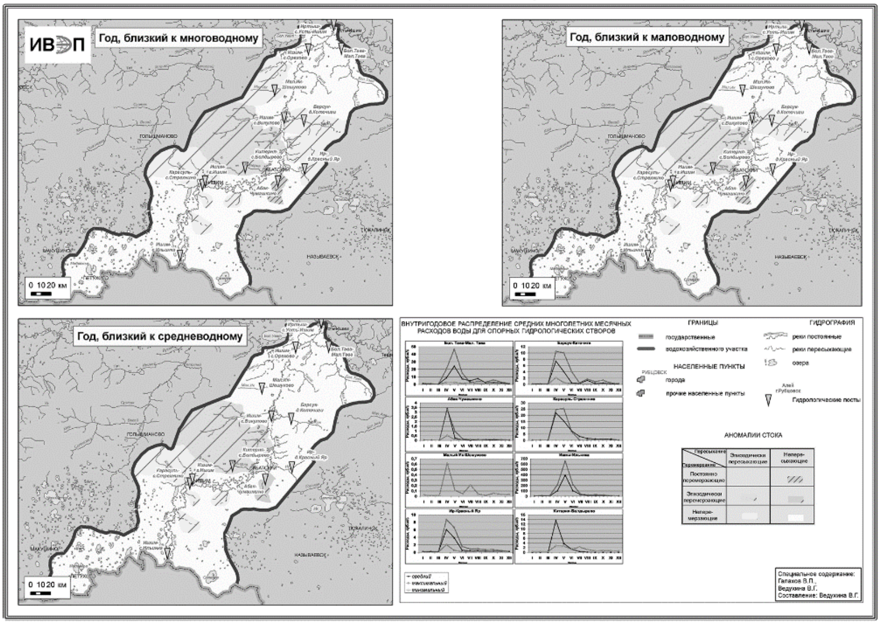


Fig. 9.3 Maps of flow anomalies (Ishim River)

Estimates of the prevalence of findings on hydrologically unexplored water bodies were performed on the basis of analogies. Water objects were grouped according to the peculiarities of the landscape structure, area and power characteristics (presence of lakes and wetlands as sources of supply).

Flow anomalies are shown on the map by the method of qualitative background (bivariate choropleth). The legend is based on the principle of the matrix and allows a comprehensive display of different combinations of the phenomena of freezing and drying up (Fig. 9.4).

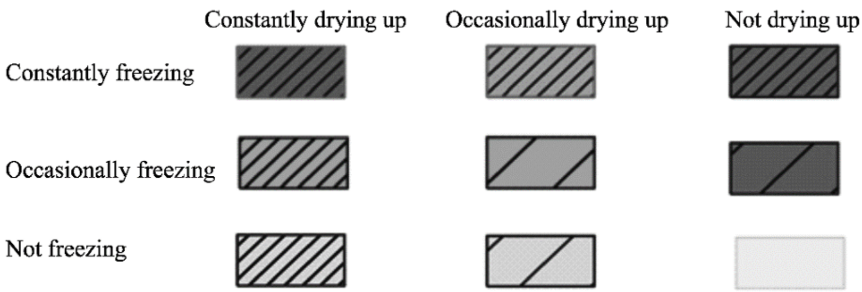


Fig. 9.4 The legend of the map of flow anomalies

Maps of industrial pollution contain data for two main interrelated characteristics: sources of surface water pollution (their localisation and quantitative characteristics) and the quality of water (Rotanova 2011b). To display sources of exposure the classification of characteristics of pollutants entering water bodies is used. According to this classification, all sources of exposure are characterised by the direct and/or indirect impact on surface water. The main sources are considered to be industrial and automobile emissions, sewage water from various enterprises; leakage of engineering structures; surface runoff from areas of industrial sites and settlements, the use of chemicals in agriculture; and the storage of industrial solid, municipal and animal waste. The map shows the volume of sewage water, sewage structure, type of land use and specific fertilizer. Volumes of sewage water are subdivided into two groups: wastewater sewage to the catchment areas, and water bodies. Thus, both direct and indirect exposure is taken into account. The volume of waste water is displayed through structural and range-graded diagrams (Fig. 9.5).

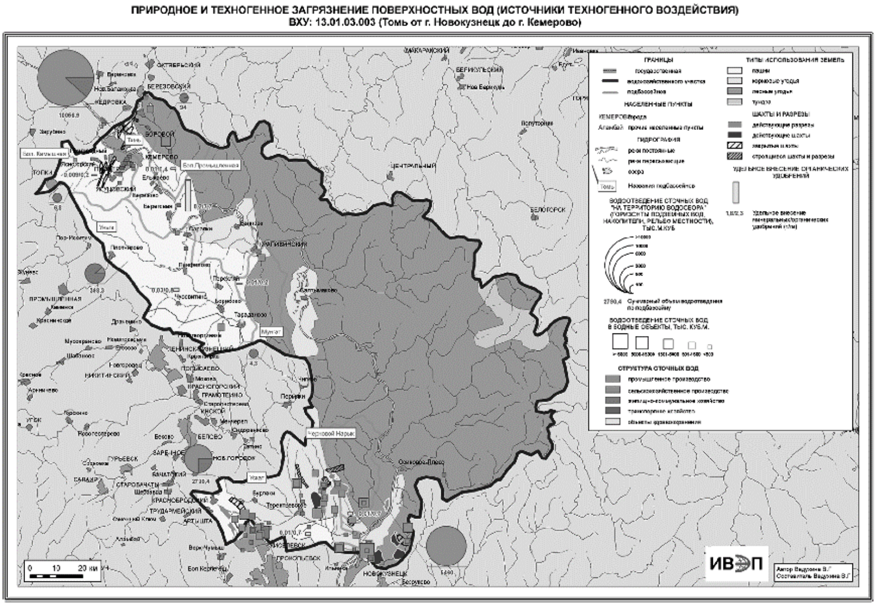


Fig. 9.5 Maps of industrial pollution (Tom’ River)

The structural division of diagrams is produced by major groups of economic activities:

- industrial production;
- agricultural production;
- housing and utilities;
- transport sector;
- other activities.

The type of land use is displayed in the qualitative background, as it describes the impact these and other types of land use have on water bodies. The basic types of land use are shown: cultivated farmland, natural grasslands, forest lands, other lands. The sources of exposure are shown by a conditional contour icon in compliance with a generalised geometry and orientation of the object being mapped.

Natural complexes (landscapes of different topological level) were used as territorial entities in performing spatial analysis. The aspects of content-richness and information value, under the different scale map models (scale 1:500 000, 1:1 000 000, 1:2 500 000), were worked through. The typification and generalisation of landscape maps for the Novosibirsk, Omsk, and Kemerovo oblasts, Altai Krai and the Republic of Altai were carried out, and electronic versions of landscape maps were made.

4 Conclusion

GIAS "Water and Ecology of Siberia" solves the GIS-related and specific tasks such as:

- cartographic assessment of state-of-the-art water resources and their use in the Ob' basin including the preparation of basic digital maps and materials to be used under development of a series of situation, evaluation and predictive maps;
- cartographic investigations of the formation processes of surface water quality and quantity including: cartographic evaluation of qualitative and quantitative indices of surface waters and their influential factors, structure development of specific databases on the basin's water objects and water object monitoring, elaboration of the structure and information filling of a geoinformation-cartographic block of the expert support system aimed at water resource management in the Ob basin.

Representative information support for the development of SCUCWO, the creation of cartographical and thematic databases, and the development of pilot GIS will contribute to the formation of the infrastructure of the basic spatial data and the justified expert system for the sustainable hydrologic functioning of the Ob basin system.

Acknowledgments The work was carried out in the framework of the SB RAS project IV.38.2.5.

References

Koshkarev AV (2009) Geo-portals and maps in the SDI's ERA. In: Proc of International Conference Intercarto/InterGIS 15 Sustainable Development of Territories, Perm (Russia) – Ghent (Belgium), p 242-246

Lovtskaya O, Rotanova I, Sutorikhin I (2007) Mathematical-cartographic support for creation of "Water and Ecology of Siberia" Geoinformation–Analytical System. Computational Technologies 12(3): 66-72

Moellering H (ed) (2006) World Spatial Metadata Standards: Scientific and Technical Descriptions, and Full Descriptions with Crosstable (The ICA Commission of Spatial Data Standards). Elsevier, Amsterdam

National Standards (2006) GOST R 52573–2006 Geographical information. Metadata, Standartinform, Moscow

Pryazhinskaya VG, Levit-Gurevich LK and Yaroshevsky DM (2011) Technical support to integrated water utilization and protection schemes. http://ru.scribd.com/fsdejournal on Jul 17, 2011

RF Government Regulation (2006) On development, approval and realization procedure of schemes of water objects integrated use and conservation, alterations to these schemes. RF Government regulation of December 30, 2006 N 883. RF Book of laws, vol 5, art. 651

RF Ministry for Natural Resources (2007a) On approval of methodical guidelines on schemes of water objects integrated use and conservation. 04.07.07 N 169

RF Ministry for Natural Resources (2007b) On approval of method for hydrographic zoning of the RF territory". Registered by RF Department of Justice on 23.05.2007, registration number 9538, Moscow

RF Water Code (2008) 74-FZ of 03.06.06 in Federal Law of 14.07.2008 №118-FZ

RFGD (2006) The concept of creation and development of RF spatial data infrastructure. Russian Federation Government Directive No. 1157-r, 21.08.2006 http://www.businesspravo.ru/Docum/DocumShow-DocumlD-114943.html

Rotanova IN, Lovtskaya OV, Vedukhina VG (2011a) Experience creating hydroecological geo-analytical system of the Ob basin. Informational Bulletin of GIS-Association 1(78): 22-26

Rotanova IN, Lovtskaya OV, Vedukhina VG (2011b) Geoinformation hydroecological analysis of surface water in the Ob basin. In: Modern Problems of Applied Mathematics and Mechanics: Theory, Experiment and Applications. International Conference, devoted to the 90th anniversary of Professor Nikolai N. Yanenko http://conf.nsc.ru/niknik-90/reportview/40806

Rotanova IN, Vedukhina VG, Cimbalei YM (2005) Cartographic analysis of water and environmental problems of the Altai Territory to optimize water use and water protection. Polzunovskiy Bulletin 4/2: 107-113

Rotanova IN, Vedukhina VG and Cimbalei YM (2009) Water-ecological mapping based on GIS. The World of Science, Culture and Education 2 (14): 23-26

Shoygu SK, Bolov VR, Komedchikov NN, Trokhina NB (eds) (2010) Atlas of natural and technological hazards and risks of emergencies in the Russian Federation. Design. Information. Cartography, Moscow

Vedukhina VG, Kuznyak YaE, Lovtskaya OV, Rotanova IN (2011) Geoinformation provision for integrated studies in the Ob'-Irtysh basin. In: Shokin YuI, Chubarov LB, Rotanova IN (eds) Geoinformation Technologies and Mathematical Models for Monitoring and Management of Ecological and Socio-economic systems, Russian Academy of Sciences, Siberian Branch, Institute for Water and Environmental Problems, Pyat' Plyus, Barnaul, pp193-200

Vedukhina VG, Rotanova IN, Lovtskaya OV (2009) The Ob'-Irtysh basin as an object of GIAS "Water and Ecology of Siberia". In: Proceedings of Conference on Water Problems of Large River Basins: Ways of Solution, Agenstvo Reklamnykh Tekhnologii, Barnaul, pp 567-579

Vinokurov YuI, Puzanov AV, Bezmaternykh DM (eds) (2012) The Present Situation of Water Resources and Water-Management Activity on the Ob'-Irtysh Basin. Izdatel'stvo SB RAS, Novosibirsk

Cartography and Geovisualisation Techniques in the Exploratory Study of Geodiversity: Almeria Case Study

Rufino Pérez-Gómez and Juan-José Ibáñez[1]

Abstract The purpose of this chapter is to describe some research into geodiversity, and conducted in the Almeria Province from a cartographic and GIS perspective. This is an area of special interest because of its aridity and severe desertification processes. Geodiversity is becoming a research topic of increasing interest among earth scientists as it may play a key role when evaluating the ecological, environmental and social value of a territory. In pursuing that goal many earth scientists use multiple mathematical techniques, such as diversity statistics and models, to end up with summary statistics which best describe the general properties of the phenomena. However, in this study, we have focused on the use of cartography, and GIS concepts and tools, in order to conduct modelling and analysis. In this study the drainage basins were chosen as basic spatial units for the analysis. Supplementary analysis of the environmental variables was conducted in relation to elevation, slope and aspect zones. The multiple and varied geoprocessing tasks enabled us to obtain many 2D and 3D map-like representations that are very useful in identifying general trends and specific patterns.

1 Introduction

Due to recent developments in data capture technologies our society is producing great volumes of digital information at a fast pace. All this data constitutes a wealth of valuable information about many aspects of the natural and socio-economic environments. We need to develop methods to properly explore, combine, analyse, visualise and communicate the findings meaningfully (Heer et al. 2010). On the other hand, cartographers are specialists in geographic knowledge

Rufino Pérez-Gómez
Universidad Politécnica de Madrid (UPM),
Madrid, Spain
Tel: +34 91 3367921. Fax: +34 913322560. e-mail: rufino.perez@upm.es

Juan-José Ibáñez
Centro de Investigaciones sobre Desertificación, CIDE (CSIC-UV), Km 405. Apdo. Oficial 46113, Valencia, Spain

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_10,

transfer. Bertin's work (1981, 1983, 2010) provided an overall theoretical framework which allowed cartographers to gain more control over the communication process by enabling them to produce clear, legible and effective maps when conveying complex spatial phenomena. Other relevant studies dealing with the visualisation of geographical space and related temporal issues have been published (Andrienko et al. 2010). In relation to this matter, the International Cartographic Association (ICA) state in its research agenda (Virrantaus et al. 2009) that one of its main commitments is *'…to ensure that geospatial information is employed to maximum effect for the benefit of science and society'*.

Biologists have been using mathematical indices and models of diversity to carry out ecological studies for many years, and to research the factors involved in the structure and organisation of ecosystems. However, interest in exploring, comparing and quantifying the diversity of other environmental factors such as soils, land forms or lithology, is much more recent. In this respect, authors such as Ibáñez et al. (1995) explained the relevance of this matter by pointing out that the characterisation and quantification of diversity of landform, rock and soil, as non-renewable natural resources, should be taken into account when estimating a territory's ecological value. The different methods used to estimate and quantify diversity may be grouped into three classes (Magurran 1988): indices of richness, abundance distribution models, and indices based on the proportional abundance of objects. At this point, it's worth mentioning findings obtained in previous studies. Ibáñez et al. (1990, 2005a) were the first to show how the evolution of fluvial systems induces an increase in the complexity of pedogeomorphological landscapes. Many ecologists have used archipelagos as natural laboratories to research the mechanisms that produce biodiversity patterns and, at the same time, to test diversity estimation techniques. Following this research strategy, Ibáñez et al. (2005a, b, 2011) conducted comparative analyses using the islands formations of the British and Aegean Archipelagos. They concluded that pedodiversity area relationships and biodiversity area relationships appear to be similar. Finally, authors such as Ibáñez and Efland (2011) are considering a "Theory of Island Pedogeography" in a similar approach to the "Theory of Island Biogeography" of MacArthur and Wilson (1967).

However, the approach in this chapter is not an in depth analysis of the diversity statistics with mathematical techniques, but rather focuses on cartography and GIS methodologies to analyse and visualise spatial patterns.

2 Materials and Methods

2.1 Almeria Case Study

The Almeria province is located in the south-east of Spain and is an area of great interest to the fields of soils and geology research, because it is one of the most arid zones in Europe. Many studies have been conducted in this area due to its particular characteristics and high degree of desertification. The LUCDEME project is a good example of these types of studies and tries to fight against desertification processes in the Mediterranean area. This project was proposed and managed by the Spanish Ministry of Agriculture, Fishing and Food (www.magrama.gob.es). The working scale for the Almeria Project is 1:100.000 and the spatial data used for the study was obtained from the following sources:

- Reference data, such as administrative units, rivers and drainage basins, together with the lithology and vegetation series thematic layers, were downloaded from the Spatial Data Infrastructure of Andalucía (www.ideandalucia.es).
- The soil vector map at scale 1:100.000 of LUCDEME project (2004) was used.
- Digital Elevation Models (DEMs) were downloaded from the National Geographic Institute website (www.ign.es).

All this data was integrated into a spatial geodatabase. For this purpose ArcGIS 10.1 software was used. Some editing and updating tasks were conducted to homogenise and prepare the data for the analysis. Afterwards, geoprocessing workflows were designed and executed, aiming to obtain the required elements for the geodiversity analysis.

2.2 Geoprocessing workflows

Following the aforementioned Magurran's classification, the proportional abundance of objects is most frequently considered when estimating diversity. From this point of view, diversity may be divided into two elements. The first relates to "*richness*". This is the number of objects present (e.g. the number of different soil types according to a certain classification). The second is "*evenness*" (a concept referring to the relative abundance of each object, i.e. the relative area occupied by each type of soil). This distinction is logical, since for two different pieces of land with the same area and identical richness, the most diverse piece will be that where the different types occupy an equal area and are therefore equally probable. Although there are several indices used in the literature, most ecologists prefer Shannon's Index for its ease of calculation (Ibáñez et al. 1995, 2013). In calculating Shannon's Index, any logarithmic base can be adopted. The natural logarithm will be used in this chapter. Its mathematical expression is as follows:

$$H' = \sum_{i=1}^{n} p_i \cdot \ln p_i, \quad (1)$$

where H' is the negative entropy (negentropy) or diversity of the population, and p_i is the proportion of individuals found in this object. In our study we calculate p_i as the percentage of surface area occupied by a particular object (soil type, lithology type, etc.) within the sample or statistical area that, in our case, is each indi-

vidual drainage basin of the project area.
Once we obtained the original database with all the reference and thematic information, edited and updated, we proceeded to isolate each individual drainage basin in the study area. Each drainage basin was assigned a *Rank* using the Horton-Strahler method which represents the stream order or its branching complexity (Ibáñez et al. 2013). These single drainage basins will behave as the sample units for diversity statistics estimation in any of the environment variables considered (lithology, soils and vegetation series). For this purpose we created three geodatabases, each one storing all the thematic polygons extracted by the individual basins. After the clipping operations, the new soil tables needed further updating. This is because the original soil polygons were cut with every basin polygon producing new smaller polygons in the outer boundary of the basin. Although the areas of the new polygons are automatically updated by the geodatabase, the 48 fields containing the surface of the soil components are static and need to be modified accordingly. To update many attribute fields in many tables, may mean several thousand of time–consuming manual editing operations, and so a model is created in order to execute the geoprocessing tasks automatically and speed up the process.

The next step was to obtain the diversity statistics through the following process:

- Each attribute table with the thematic polygons by basin was exported to a spreadsheet.
- All the objects within the basin were sorted by the right attribute (Vegetation type, lithology type, etc.).
- The total area for each object class was obtained.
- The proportional abundance of each object type, within each basin, was computed as the ratio between the sum of the areas of all polygons of the same type and the particular drainage basin area.
- The richness and Shannon’s Index were calculated.

After this process, we obtained richness and Shannon’s Index values for all drainage basins and the three environmental variables (lithology, soils and vegetation series). Any individual basin now had six new individual numeric attributes to work with and analyse.

3 Geomatic Analysis

The geomatic analysis, which follows, encompasses three types of work: thematic cartography, 3D visualisation techniques and zonal statistical analysis with raster techniques. In this study we consider two abiotic factors, lithology and soils, together with a vegetation series to allow experts to compare the geodiversity values with biodiversity values in the area. From the cartographic point of view, however, all variables, as spatial distributions, were treated in the same way.

3.1 Thematic cartography

All drainage basins in the Almeria province map were assigned a rank using the Horton-Strahler method, reflecting their branching complexity or stream order, and richness and Shannon's Index values for all the drainage basins were obtained during the mathematical calculations. All this information, generated in a spreadsheet, was integrated back into the spatial database as numeric attributes to enable the map production workflows. At this point, it is important to remember the need for sound cartographic principles in the map design and production phases, as reflected upon by Morita (2011) and Ormeling (2013) in their papers.

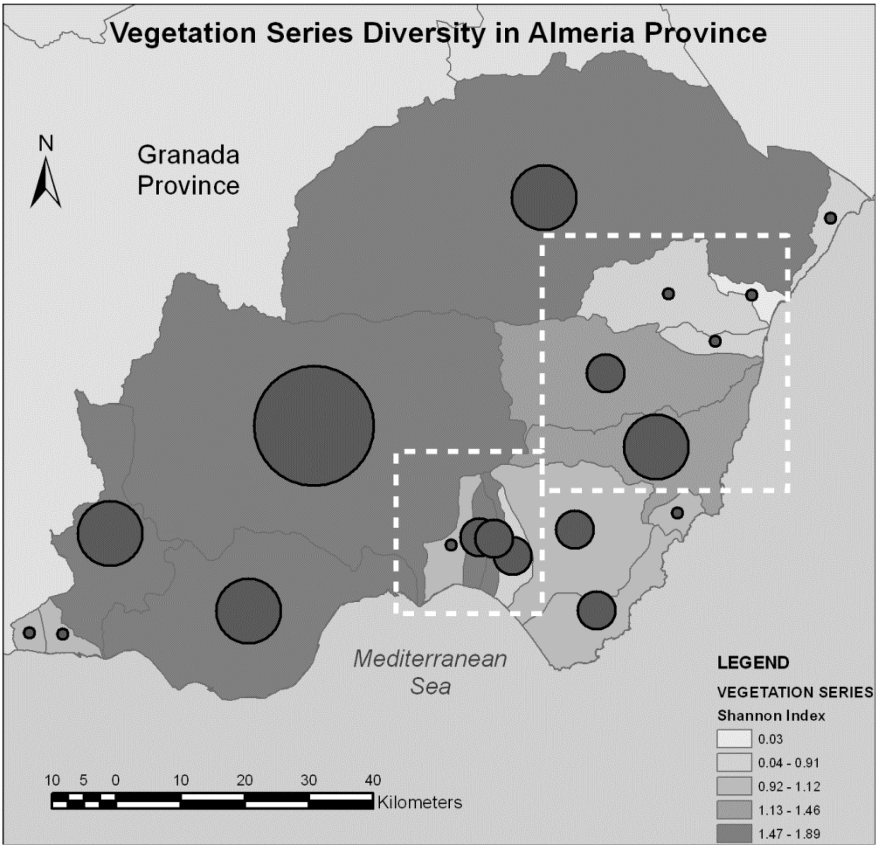


Fig. 10.1 Vegetation Richness and Shannon's Index values by drainage basin in Almeria.

Figure 10.1 represents the diversity values of the vegetation series layer. This map shows the richness and Shannon's Index variation by drainage basin in the Almeria province. Using the map we can draw conclusions about the mapped area:

- As a general rule richness and Shannon's Index increase with the rank and area of the basin.
- The northeast of the Almeria province presents low vegetation diversity values.
- Some local variations in richness and Shannon's Index values have been observed among neighbouring basins of the same rank. The bottom square illustrates this point with four Rank 1 basins. Similarly, the top right square contains three Rank 2 basins with significant vegetation diversity values.

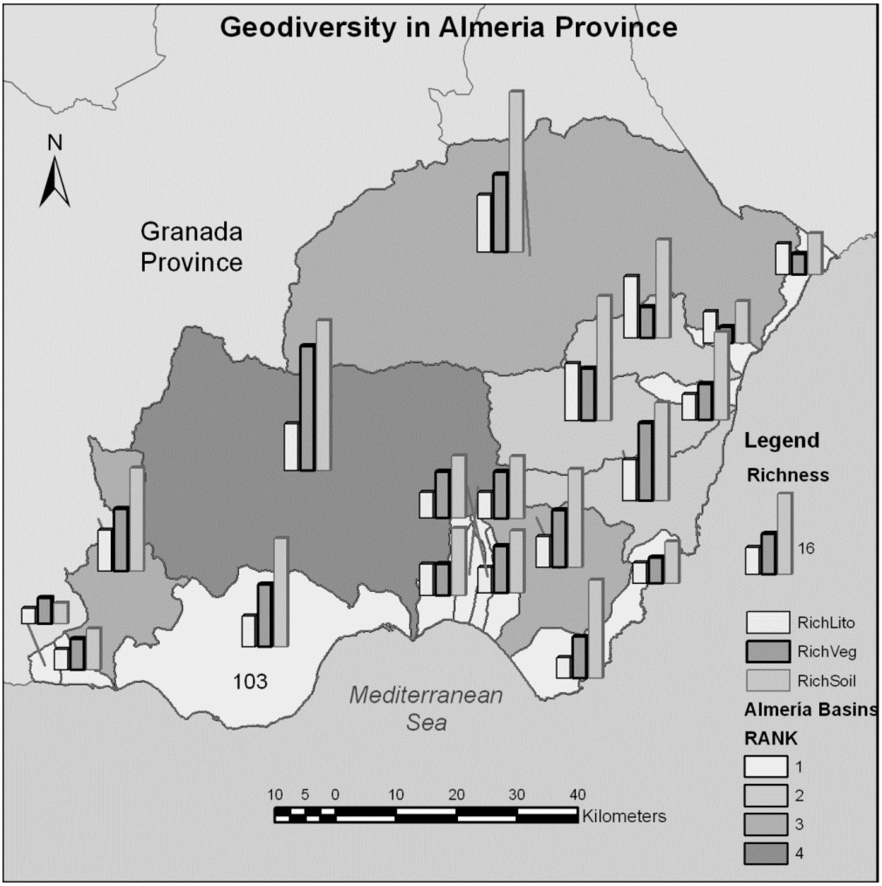


Fig. 10.2 Variation of Richness values with the basin rank/area in Almeria province

In the same manner, similar maps have been produced to analyse lithology and soil diversity by drainage basin. We can also compare the diversity values (e.g. richness values) of several environmental variables at the same time and in relation to the rank of the basin with a polyanalytic map. This strategy is used in Figure 10.2. The map shows the three richness values, in different colours, with their associated proportional symbols. At the same time, the basin rank layer provides a suitable background to analyse variation in the diversity statistics. The reading and interpretation of this map allows us to make some comments:

- As a general trend, richness grows with the rank/area of the basin. Low rank basins have smaller areas because they are related to small rivers which have their origin near the coastline.
- Soil seems to have the highest richness values, while lithology the lowest in most basins. This may reflect particular spatial properties in the area, but it may also be caused by differences in the degree of generalisation of the input maps.
- In the NE part of the province a decrease in vegetation richness is observed.
- Basin 103 ("*Campo de Dalías*") is a special area where many glasshouses have been built. It actually represents the combination of several Rank 1 basins. In this particular case the area is more relevant than the rank for the mathematical analysis.

A careful reading and interpretation of each map, as made in the two previous examples, enables us to identify general trends, zonal patterns and particular cases to which the earth scientist experts should be able to attach adequate "geographical meaning".

3.2 3D Visualisation Techniques

In addition to the traditional 2D mapping techniques, we can also explore and analyse any geographic phenomenon by building 3D models. We then identify general trends, zonal patterns and particular cases as we do in the map reading and interpretation phase. Figure 10.3 illustrates this technique with a "3D diversity model" applied to the soil layer. In this model the Shannon's Index is represented by a colour ramp or value scale in which darker colours are assigned to the highest diversity values. "Richness values" are shown by means of extrusions or vertical extensions and, by doing so, 2D polygons turn into solids or 3D blocks. Once the models are created we can work with them in different ways:

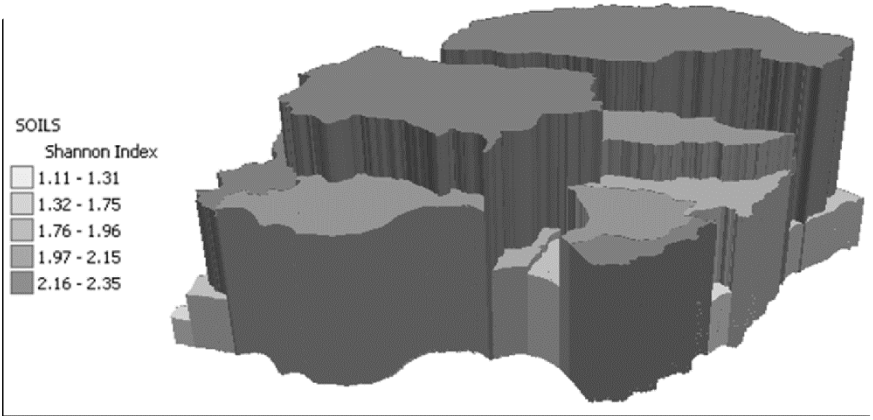


Fig. 10.3 3D soil diversity model in the Almeria province

- Analyse each model separately to identify global and zonal patterns and local or particular cases.
- Compare several models simultaneously to identify the different spatial properties of the basins.
- Create and store perspective or 3D views that best communicate the previous findings. Create animations that show the results and most relevant findings in a "dynamic" manner. Alternatively, these animations can be exported to video for exchange with other users.

3.3 Zonal statistical analysis with raster techniques

This geodiversity study was conducted using the drainage basins as the spatial or sample unit of analysis. Another interesting approach would be to observe and compute the diversity statistics variation with the components of the 3D environment. After reclassifying the continuous surfaces we obtained three raster surfaces with elevation zones, slope zones and aspect zones. The three vector layers that describe the environmental variables were converted to raster format (VRC). Once we had the zone layers and the thematic layers in raster format (also called "value layers") we could conduct a zonal statistical analysis. Instead of a new raster, the zonal statistic function produces a table of statistics and an optional graph or chart. Figure 10.4 shows a graph of this type which reflects a clear decrease in the number of vegetation types ("*richness*") with the elevation. This is particularly clear above the altitude of 1000 meters. In similar fashion we analysed the variation of vegetation according to the slope zones and aspect zones. The results show how the number of vegetation types change slightly with the slope and that there is no change with the orientation zones.

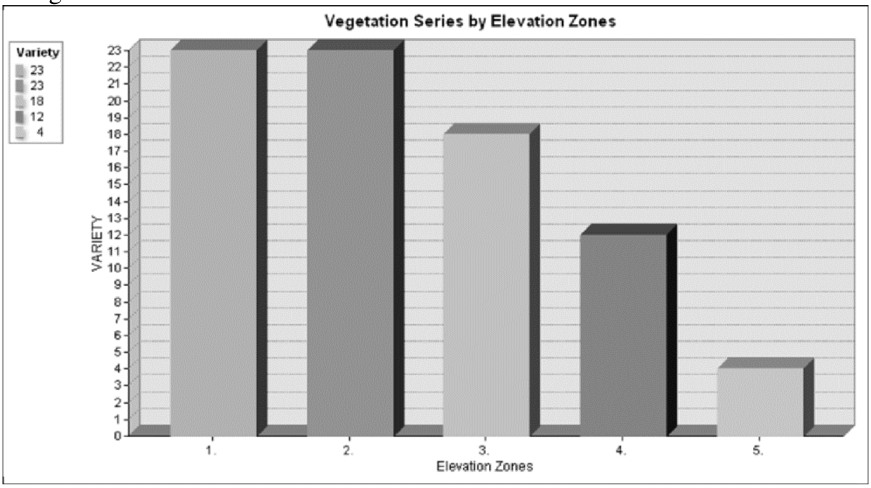


Fig. 10.4 Change in vegetation series variety with elevation zones in Almeria Province

The procedure was repeated by combining the lithology and soil thematic layers with the three described zone layers in order to study their particular spatial properties. The main findings are:

- A clear decrease in the number of lithological types is observed with the elevation zones. The number varies from 15 lithological types in the first zone (0-500m) to only 5 types in the last zone (over 2000m). However, no significant variation is observed with the change in slope and the change in orientation.
- The number of soil types shows a clear decrease when the elevation increases. A slight decrease is also shown in the number of soil types with the slope zones, and no change with the aspect zones.

4 Conclusions

To analyse the landscape characteristics soil and earth scientists use multiple diversity statistics and graphics. This chapter describes how the study of any geographical phenomenon with cartographic and GIS technologies allows us to explore, compare, assess and quantify the different spatial patterns embedded in the geographic distributions of any study area or environment. Some conclusions from this cartographic and GIS approaches are:

- Analytical maps are useful tools to identify and show global trends, zonal patterns (e.g. high or low unusual values), local variation zones and local or particular cases of any spatial distribution (e.g. soil diversity map).
- Polyanalytic maps enable the simultaneous representation of different variables at the same time. This allows the analysis and comparison of several spatial

properties to identify relationships or recognise areas of special characteristics (e.g. three variable richness values on top a basin rank/area layer).

- 3D models are also a very visual and attractive manner to reveal the structure and main patterns of a geographic distribution.
- The zonal statistical analysis offers interesting features such as the study of richness variation with elevation, slope and orientation zones. (e.g. soil diversity variation with elevation or slope zones).
- Because of the previous conclusions, the thematic maps, 3D models and zonal statistical analysis are useful alternative manners for analysing the richness-area relationships and to conduct geodiversity studies.
- The results of the geodiversity studies might be used in environmental management and assessment activities, the design of natural reserves or geoparks or the elaboration of environmental legislation.

In short, we have tried to illustrate how maps and Geographic Information products contribute to the process of knowledge acquisition and are very important elements supporting scientific investigations.

References

Andrienko G, Andrienko N, Demsar U, Dransch D, Dykes J, Fabrikant SI, Jern M, Kraak MJ, Shumann H, Tominski C (2010) Space, time and visual analytics. International Journal of Geographical Information Science 24:10, 1577-1600

Bertin J (1981) Graphics and Graphic Information Processing. De Gruyter, Berlin/New York

Bertin J (1983) Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin Press, Madison, WI

Bertin J (2010) Semiology of Graphics: Diagrams Networks Maps. ESRI Press, Redlands, CA

Heer J, Bostock M, Ogievetsky V (2010) A tour through the visualisation zoo. Communications of the ACM, 53: 6, 59-67

Ibáñez JJ, Jiménez-Ballesta R, García-Álvarez A (1990) Soil landscapes and drainage basins in Mediterranean mountain areas. Catena 17, 573-583

Ibáñez JJ, De-Alba S, Bermúdez FF, García-Álvarez A (1995) Pedodiversity: concepts and measures. Catena 24, 215-232

Ibáñez JJ, Caniego FJ, San-José F, Carrera C (2005a) Pedodiversity–area relationships for islands. Ecological Modelling 182:3-4, 257-269

Ibáñez JJ, Caniego FJ, García-Álvarez A (2005b) Nested subset analysis and taxa range size distributions of pedological assemblages: implications for biodiversity studies. Ecological Modeling. 182:3-4, 239-256

Ibáñez JJ, Effland WR (2011) Toward a Theory of Island Pedogeography: Testing the driving forces for pedological assemblages in archipelagos of different origins. Geomorphology 135, 215-223

Ibáñez JJ, Vargas RJ, Vázquez-Hoehne A (2013) Pedodiversity State of the Art and Future Challenges. In: Ibáñez JJ, Bockheim JG (ed) Pedodiversity, CRC Press (Taylor and Francis Group) Boca Ratón. California., pp. 1-28

MacArthur RH, Wilson EO (1967) The Theory of Island Biogeography. Princeton University Press, Princeton, NJ

Magurran AE (1988) Ecological Diversity and its Measurement. Croom Helm, London

Morita T (2011) Reflections on the works of Jacques Bertin: From sign theory to cartographic discourse. The Cartographic Journal 48: 2 pp. 86-91 Special Issue May 2011

Ormeling F (2013) Cartography as a tool for supporting geospatial decisions. GIM International, 27: 8

Virrantaus K, Fairbairn D, Kraak MJ (2009) ICA research agenda on cartography and GI science. The Cartographic Journal, 46: 2, 63-75

LUCDEME Project website (2013)

http://www.magrama.gob.es/es/biodiversidad/temas/desertificacion-y-restauracion-forestal/lucha-contra-la-desertificacion/lch_lucdeme.aspx. Accessed 10 December 2013

The Geographical Size of a Country and Methods for Its Calculation and Mapping

Vladimir S. Tikunov, Olga Yu. Chereshnya[1]

Abstract There have been many attempts to systematise elements that form the power of the state, and bring them into one formula of power. Military experts have been working on the invention of such formulas since the 1950s. Many scientists have noted the unreliability of the results of previous national power calculations. Their doubts are caused by the complexity of calculating factors and indicators such as the “will to pursue national strategy”. One of the indices for assessing the power and potential of a country is the general geographic size index (GSI). GSI includes only three components: area, population and economic strength, without considering political power and culture, which are difficult to assess. These indicators may seem primitive for calculating state power, but they have several important advantages: there is data for almost all countries, they do not require expert assessment and they are reliable. GSI clearly illustrates changes in the world power balance and the dynamics of countries’ development.

1 A Country's Geographical Size

What is the size of a country? Most people consider only surface area, but when people think about the size of a country, it seems to involve not only an area, but also political and economic power. For example, few people in Europe know that the Ukraine is the biggest country in Europe. This is due to its low role in the political arena and a weak economy, compared with other European countries. There are other big countries, the size of which is not in doubt. They have powerful economics, large areas, big populations, strong politics and culture. Science cannot be based on narrow-minded perception, however, it needs accurate data.

There have been many attempts to systematise the elements that form the power of a state, and create a formula for this power (Tikunov 2009). Military experts have been working on the invention of such formulas since the 1950s (Balakhontsy and Kondratiev 2010). An example is the "national power equations" of Ray Cline. He thought that the power of the state was “a set of material

V. S. Tikunov, O. Yu. Chereshnya
M.V. Lomonosov Moscow State University
Geographical Faculty
Moscow 119991 Russia
vstikunov@yandex.ru, cherry.purplerain@gmail.com

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_11,

and spiritual powers of the country and the ability to mobilise them in war, and that it consisted of economic, social, technological, political and military capacities". In his book "The Power of Nations in the 1990s: A Strategic Assessment" he created a formula that describes territorial size, population, economic capability, and military power for strategic purposes and national will. He calculated the power of a state mathematically using the equation: Pp = (C + E + M) x (S + W) where, Pp is the "perceived power of a state, C is the critical mass which includes territory and population, E is the economic power, M is the military power, S is the strategic purpose, and W is the will to pursue national strategy.

For example, it rated the US as 35 and the erstwhile USSR as 67.5. Cline (2002) also describes power as "a subjective factor" and uses the term "perceived power" in his formula. He suggests that "real power" is something different from "perceived power" but does not clarify this distinction.

There are many other approaches to calculation. Comprehensive National Power (CNP) developed was by Chin-Lung Chang and the Composite Index of National Capability (CINC) is a statistical measure of national power created by J. David Singer and others). Many scientists have noted the unreliability of the results of national power calculations. Their doubts are caused by the complexity of calculating factors and indicators such as the "will to pursue national strategy". Such indicators are often based on expert judgment rather than on accurate data. Therefore, it is necessary to develop a simple and reliable way to compare countries.

One of the indices for assessing the power and potential of a country is the general **geographic size index** (GSI). GSI includes only three components: area, population and economic strength, without considering the political power or culture, which are difficult to assess. These three basic geographic indicators are like the classic factors of production: land, labour and capital (Tikunov and Treyvish 2006). These indicators may seem primitive for calculating state power, but they have several important advantages: there is data for almost all countries, they do not require expert assessment and they are reliable.

Territory is the most important resource for any country, it is an arena for human activity, the source of minerals and land for agriculture. Large-area resource-poor countries do not exist. Vast areas bring as many problems as they give benefits, however (Hill and Gaddy, 2003) These problems are described in post-Soviet geography. The concept of small countries was described by Zimin (1990) and later by Gorkin A and Lipets Yu (2003). A large area entails the problems of internal relations, and national, religious and cultural conflicts.

Population involves the people who are workers and consumers (Ratzel 1897). A populous country includes strong social contrasts: even if it is richer on average, there will also be more poor people. Of course, the population is heterogeneous, and in different countries, the same number of people may have a completely different potential. For example, the population of Singapore and Turkmenistan are almost identical, but there is an obvious difference in their productivity, in their level of technology and its potential. It might thus be better to evaluate the human capital of the country, rather than the population. Unfortunately, at this stage, human capital calculation methods are imperfect and cannot provide objective results.

Finally in today's world economic power is objectively a key indicator showing the power of a country (Treivish 2005). GSI is one of the most common ways used for calculating a country's economy - gross domestic product converted to international 2005 year American dollars using purchasing power parity rates.

One method of calculation is following: instead of the average size index (arithmetic average of the percentage of a country's shares in the world by territory, population and GDP) it uses the evaluative algorithm. It includes normalisation of initial indicators by the formula:

$$\hat{x}_{ij} = \frac{|x_{ij} - {}^{0}x_j|}{|{}_{max/min}x_j - {}^{0}x_j|},$$

$$i = 1,2,3,\dots,n;$$
$$j = 1,2,3,\dots,m;$$

where ${}^{0}x$ is theoretically possible worst value, here is zero for all three indicators (as ${}^{0}x$, you can also use the worst-case values for each indicator occurring over the whole period in time);${}_{max/min}x$ is the most different from the ${}^{0}x$ values of parameters, n is the number of territorial units, m is number of indicators used for the calculations. Ranking is carried out by comparing all territorial units on a conditional basis, characterised by values of ${}^{0}x$. This is done using the Euclidean distance (d°) as a measure of proximity of all territorial units to a conditional basis (a worst-case values ${}^{0}x$ throughout a range of indicators). Processing of the array using principal component analysis for the purpose of orthogonalisation and a "convolution" system of indicators was then used. In the experiment, all the original indicators have equal weight, although it was assumed that weight meaning can be different and can change through history. To apply the weights we need a reliable basis, however and today there is no such basis (Tikunov and Treyvish, 2006).

2 Size of Historical Countries in the Twentieth Century

Global history labels the year 1900 as one of Pax Britannica. The power of the British Empire was expressed in maximal power (Mackinder 1904), being inferior to China by only one per cent of the world's population. The then weak Chinese Empire was barely in control of some of its regions and could not be a rival to Britain. Instead the USA pursued the British colossus economically, which included the poorest parts of the world. The Russian Empire dropped twice as far behind the leading GSI, and France and Germany, with their colonies, by three to fivefold. These six states were the giants of the early 20th century.

Scoring on GSI one can mark out other groups of states. Brazil, the Netherlands, Austro-Hungary, Turkey and Japan could be considered large states. Medium-sized powers such as Persia/Iran, Sweden, Ethiopia, Peru and Siam/ Thailand might have indices two to four levels lower. The next grade is represented by smaller countries such as Serbia, Greece and Ecuador etc., and the very smallest were the states of Central America, Liberia, and Luxembourg with Montenegro as the last member in the list for 1900.

In 1925, the British world stood fast and had even expanded as Britain took the German colonies in Africa, but it yielded to the USA in terms of GDP. China had lost Mongolia, Tibet (de facto) and Manchuria (partly) but remained second in size. Russia, now the USSR, was fourth. Germany held the sixth place by GSI, but lagged notably behind the fifth place France and left the group of giant states. The Austro-Hungarian Empire had collapsed, and Turkey diminished to the level of medium countries, augmented by new 'successor" states such as Czechoslovakia and Yugoslavia. (Maddison 1999) There were more small and smallest states, none of them very big.

The year 1950 was the epoch of two super-powers (Hooson 1964) in spite of the fact that the USSR had not overtaken China by GSI. The USA became the leader in the post-war world. Britain's empire was disappearing faster than the French; she was level with the jewel in the crown, namely India. The giant states were again six in total, with India in the place of Germany. Canada, Australia, Pakistan and Indonesia had joined the circle of large countries, the quantity of medium and small began to grow, but the smallest did not change in number.

In the 1950-75 period the number of countries in the world, above all in Africa, increased quickly at the expense of medium and especially small, which doubled, while the very smallest tripled. Although the areas of the giants were stable, China passed the USA in GSI, and the latter index was even higher in the USSR. Decolonisation removed Britain and France from the group of giants while India, Brazil and Canada moved up the ranks. Economically successful Japan and West Germany, as well as Indonesia, Mexico and Italy, grew in size.

In 2000 China widened the gap between itself and the United States; in the absence of the USSR, these two remained the “super giants”, India came nearer to Russia and Brazil was almost as large. The number of large and medium states changed to a lesser degree, while all the growth came from the small and smallest countries, the common share of which exceeded 60 per cent. The proportion of GSI by group was naturally different. Giant states in 2000 comprised 42 per cent of the world’s total, despite a constant reduction of the group's share during the century. The share of large countries has remained at the level of 23% since 1950. The total GSI of the medium sized countries exceeded 25 per cent, showing rapid growth along with the two smaller groups, although their relative weight did not reach 10 per cent.

Thus the absolute domination of the giant states sank into oblivion, but they have not left the arena. Among them China and India are the most populous but poor, the USA is the richest, and Russia has the greatest extent. Colonial empires have gone and ceded to their own former possessions such as India, Canada, Indonesia and Australia. (Turovsky 1999) In 1900 there was only one giant power in the southern hemisphere, and in 2000 there were three. China has retained the parameters of a "super giant", in no less than second position by GSI. Nonetheless, it did give up its position in the mid-20th century, when the USA and the USSR were gaining weight and Britain lost it (Tikunov and Treyvish 2008).

Many smaller states were shaking for various reasons, including the loss of territory, either early as in the case of Turkey, which grew afterwards, or later, like Belgium deprived of its Congo. Even if the area changed little, the demo-economic dimension and GSI increased or decreased.

3 GSI from 1990 to 2011

The following study was conducted for the period from 1990 to 2011, which allows us to trace the modern dynamics of GSI in stable contemporary borders. It includes most of the world (167 countries), excepting those countries for which there are no reliable statistics. This period is characterised by the fact that the process of formation of new states has slowed significantly since the mid-20th century. Recent significant changes have occurred since the collapse of the last "empire", the Soviet Union, in 1991. It allows us to count GSI in permanent borders, because all future changes in the territory of countries are minor and do not play a significant role in the final index.

The results are clearly consistent with the concept of leadership in the world arena. GSI leaders consist of all the fast-growing BRIC countries (Brazil, Russia, India and China), the world's economic leaders in GDP (US, China, Japan and Germany (List of the World Bank (1990-2012), World Bank Open Data 2013), the most populous countries (China, India, USA, Indonesia, Brazil), and those with the largest area (Russia, Canada, China, USA, Brazil). The first ten by GSI are undisputed giants of countries that we can say are large in every way (Fig. 11.1).

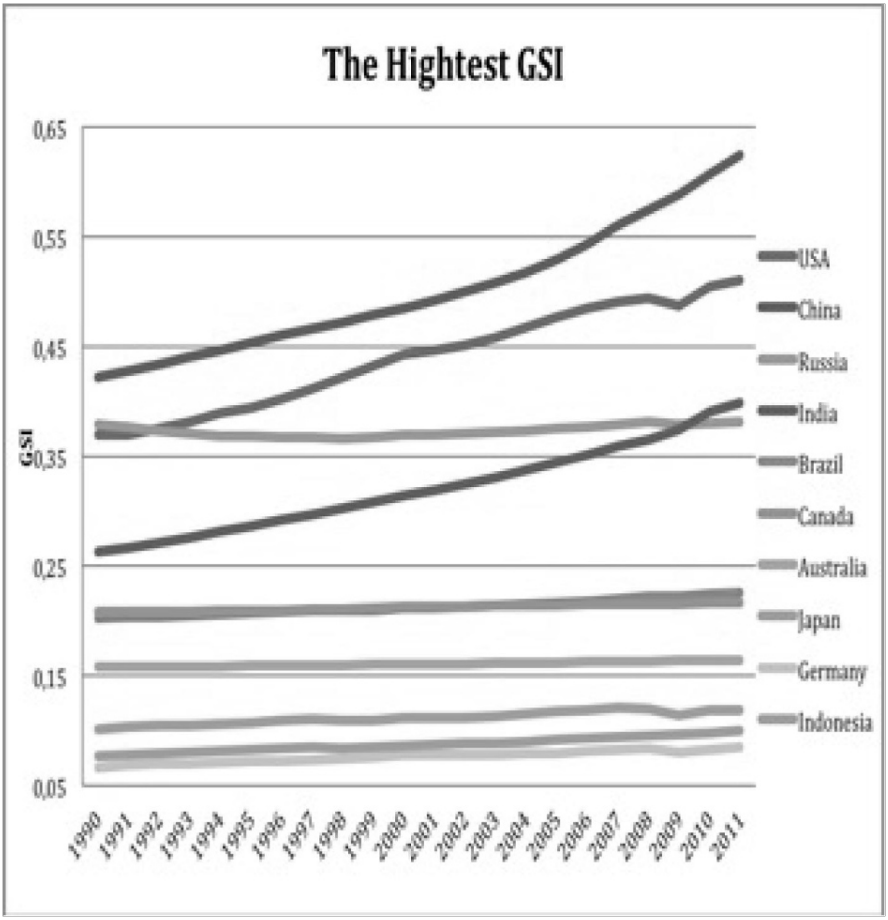


Fig. 11.1 GSI of the top ten countries, 1990-2011

China is in the first place. Stable high population growth and rapidly growing economy, plus a huge area, have allowed China to hold first place for two decades. From 1990 to 2011 its GSI increased by 0.2.

In second place, according to rates of growth in the index, is India. Its GSI grew by 0.14 since 1990. It has an even faster growing population than China, but lower GDP growth (Fig. 11.2, 11.3). This growth allowed India to move ahead of Russia in 2009 despite it having the biggest territory.

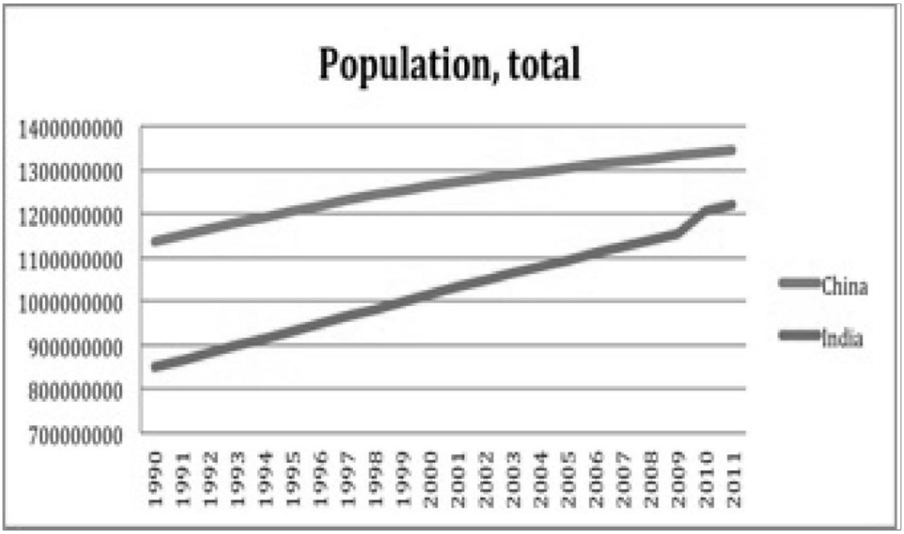


Fig. 11.2 Population of China and India, 1990-2011

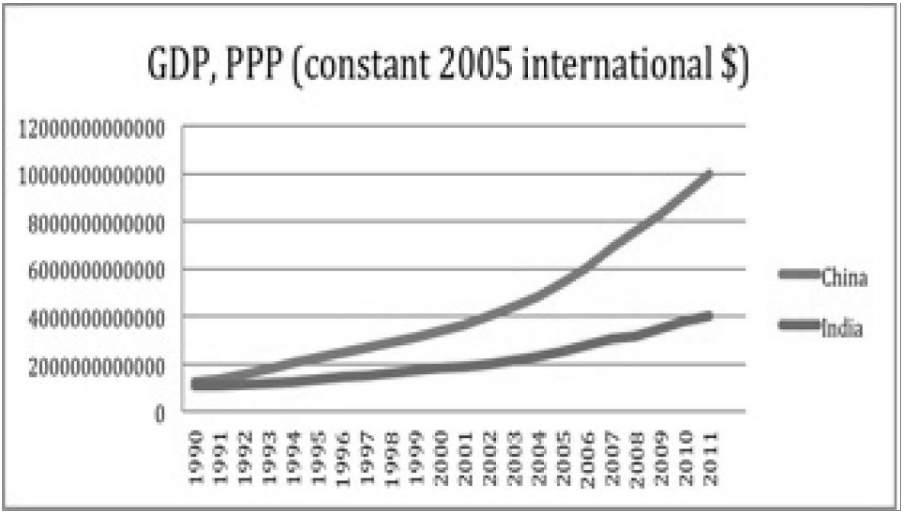


Fig. 11.3 GDP by PPP (international dollars, 2005 of China and India (1990-2011)

The GSI graph of the USA is not such a straight line. This is due to the fact that, unlike China and India where growth in GSI is part of both GDP and population growth, the USA index increased almost entirely due to its GDP, and its dynamics are strongly influenced by economic crises.

To understand the dynamics and transformation of GSI we need to understand which parameter has the greatest effect on the total index and how this influence varies with time. For this we can use RGB maps. Colour shows the ratio of the territory (red), population (green) and GDP at PPP by 2005 $ (blue) in 1990 (Fig. 11.4) and 2010 (Fig. 11.5) years.

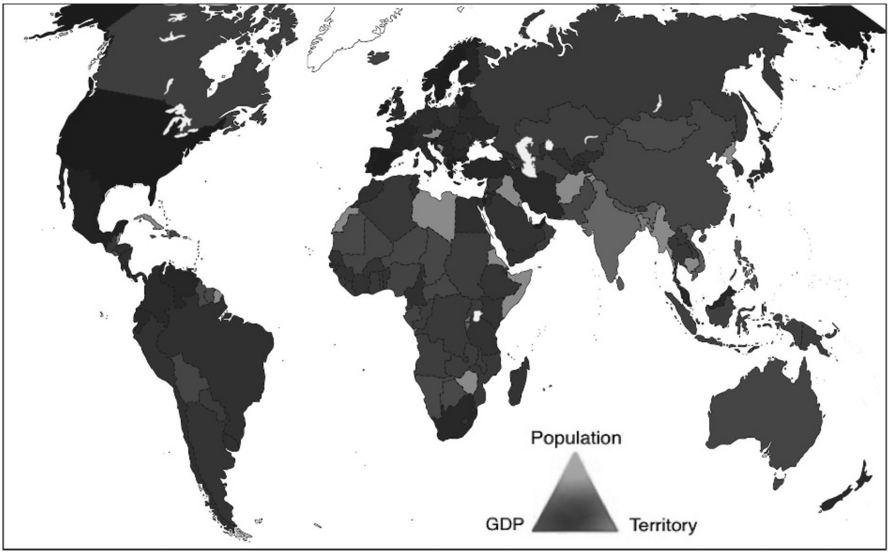


Fig.11.1 GSI of countries in 1990

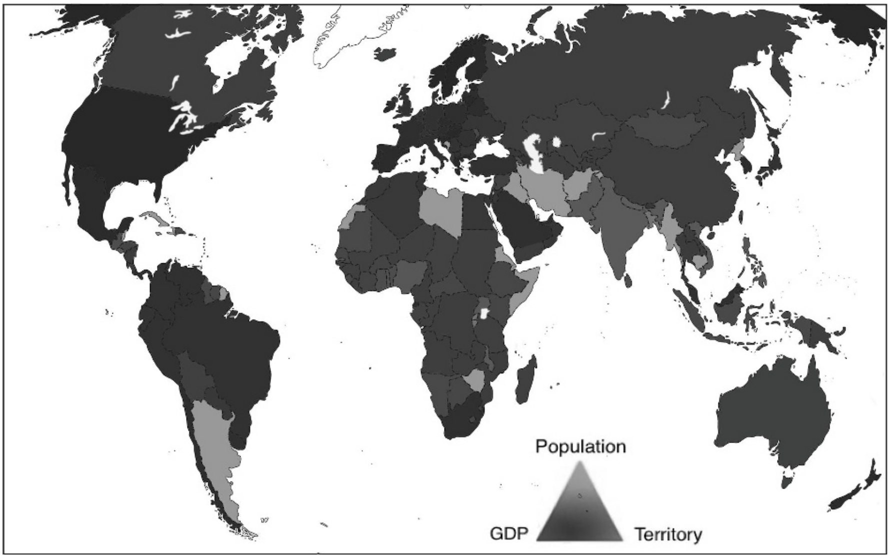


Fig. 11.2 GSI of countries in 2010

As can be seen in the map, in the index of most countries, both large and medium, a share of territory prevails, but it's percentage decreases with time. We can identify different colour groups of countries on the map. In Western Europe, the main role is played by GDP, its influence gradually decreases from west to east Europe. South-East Asia has a high index mainly due to population, despite the highest GDP and large areas of China and India. High-tech countries South Korea and Japan stand out significantly among them. Countries with a vast territory and disproportionately low GDP and population, such as Canada, Brazil and Russia, are painted red. Scandinavian countries form a separate group. They have large GDP and a large area with a small population – this is a rare case. The USA is also separate, although it has the fourth largest area in the world and third largest population, its index consists mainly of GDP.

It may be noted that significant changes have occurred in the percentage of components in the index since 1990. First of all, there is an increase of share of GDP in the countries of Southeast Asia. The economies of China, Malaysia, and Indonesia are growing more rapidly than the population. In 1990, South Korea's share of GDP and the population was almost equal in the GSI. By 2010, the share of GDP was greatly increased and now constitutes a major part of the index, along with Japan. The proportion of GDP in Latin America increased less significantly. The role of GDP had grown in Central Europe and Eastern Europe in comparison with 1990, although it still decreases from west to east. In Africa, the situation is different, there the share of GDP is almost constantly very low, but the proportion of population significantly increased in many countries.

4 Summary

In conclusion we would like to note that even such a simple index based on basic indicators, demonstrates the real picture of power very well, and illustrates the dynamics of development. The results we see on the GSI graphs and maps may all be common knowledge, but the simplicity of this approach and the possibility of a clear visualisation of countries' development processes shows the need for further study and improvement of the methods.

References

Balakhontsy N, Kondratiev A (2010) Foreign countries' capacity assessment methods, Foreign Military Review, № 11. Moscow: Agency "Voeninform", pp 101 -104

Cline RS (2002) The Power of Nations in the 1990s: A Strategic Assessment. University Press of America

Gorkin A, Lipets Yu (2003) Allocation of production in market environment. Moscow: Alfa-M Publishing House (in Russian)

Hill F, Gaddy C (2003) The Siberian Curse: How Communist Planners Left Russia Out in the Cold. Washington, D.C.: Brookings Institution Press

Hooson D (1964) New Soviet Heartland. London: D. Van Nostrand Company

Mackinder H (1904) The geographical pivot of history, Geographical Journal, vol 23 pp 421-437

Maddison A (1999) Monitoring the World Economy 1820-1992. Paris, OECD

Ratzel F (1897) PolitischeGeographie, oder die Geographie der Staaten, der Verkehrs, und der Krieges. Munich-Berlin: R. Oldenbourg

Tikunov A (2009) Integral Indicators of Spatial Models of Development of the World. Moscow: Librokom Publishing House (in Russian)

Tikunov VS, Treyvish AI (2006) An attempt to estimate geographical size of countries and their regions, Moscow University Herald, ser 5 Geography, № 1, pp 40-49 (in Russian)

Tikunov VS, Treyvish AI (2008) The geographical dimensions of countries and regions in twentieth and twenty-first centuries, Geography, Environment, Sustainability, №1,vol 1 pp 6-21

Treivish AI (2005) New Russian heartland: The demographic and economic dimension, Eurasian Geography and Economics, № 2, vol 46, pp 123-156

Turovsky RF (1999) Political Geography. Moscow-Smolensk: Smolensk Humanitarian University

World Bank Open Data (2013) http://data.worldbank.org/. Accessed 12 Dec 2013

Zimin BN (1990) Small countries as territories of high economic openness, Economic-Geographical Problems of Forming Regional Export-Import activity. Novosibirsk, pp 41-56 (in Russian)

The Academic Atlas of Czech History

Tomáš Janata, Jiří Cajthaml, Pavel Seemann and Růžena Zimová[1]

Abstract Processing historical maps for historical atlas work involves a series of time consuming and technically demanding activities. The creation of such maps is quite specific. Thematic atlases include mainly cartographic presentations of statistical (or other) data from a particular field. In historical maps not only the physical-geographical or general topographic content relative to the specific historical period must be shown, but also a historical event or situation in many cases linked to the present state. Thematic elements have multiple attributes and may reflect the quality, quantity and time of the displayed phenomenon, which must be suitably visualised. Combining cartography and geoinformatics supports the use of GIS products in the GIS implementation of a comprehensive project of a thematic cartographic atlas. This chapter summarises the basic information about the upcoming Academic Atlas of Czech History and the experience of the team of cartographers in making historic maps for the atlas – from the preparation of the atlas concept through materials used, creation of a data model, map key and particular data sets to issues related to technical aspects of the work or to balancing needs of historians and the map space delimited.

Tomáš Janata
Czech Technical University in Prague, Department of Geomatics,
Thákurova 7, Prague
e-mail: tomas.janata@fsv.cvut.cz

Jiří Cajthaml
Czech Technical University in Prague, Department of Geomatics,
Thákurova 7, Prague
e-mail: jiri.cajthaml@fsv.cvut.cz

Pavel Seemann
Czech Technical University in Prague, Department of Geomatics,
Thákurova 7, Prague
e-mail: pavel.seemann@fsv.cvut.cz

Růžena Zimová
Czech Technical University in Prague, Department of Geomatics,
Thákurova 7, Prague
e-mail: ruzena.zimova@fsv.cvut.cz

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_12,

1 Cartographic Production on the Theme of Czech or Czechoslovak History

Among the first thematic maps focusing on the history of the Czech lands are the maps of teacher and spiritual leader Aleš Pařízek from 1781 (Semotanová 2003). The first historiographical atlas, titled "The Old World Atlas" was prepared by Václav Merklas in 1850. A historical map of Czechia by František Palacký and Josef Kalousek, first issued in 1876, has become a well-known cartographic work. In the first half of the 20th century Antonín Balcar, Jaroslav Vlach, František Kameníček, Jan Macháček, Josef Brunclík, Otakar Dorazil and Jaroslav Lameš excelled among authors of historical atlases. In the period after World War Two, mainly historical school atlases were published; the "Pocket Atlas of World History" appeared in several editions in the 1970s and the 1980s.

To the 1960s in the 20th century fell the edition of the biggest thematic atlas focused on history – the Atlas of the Czechoslovak history (Purš et al. 1965). The atlas consisted of 45 sheets (approximately 86 by 50 cm) with more than 400 maps and graphs, created by the team of the Institute of Czechoslovak and World History of the Czechoslovak Academy of Sciences. Maps showing the history of national territory were also included in large (national) atlases issued in the 20th century (Atlas of the Czechoslovak Republic in 1935, Atlas of Czechoslovakia in 1966, Atlas of Slovak Republic in 1982).

At the end of the 20th century, new historical school atlases were released due to the changes in the interpretation of historical development in Czechoslovakia. In 2007 came the Otto's historical atlas by Eva Semotanová. Historical development of the area was also partly presented in the Landscape Atlas of the Slovak Republic (2002) (maps are also available in the internet application) and the Landscape Atlas of the Czech Republic (2010). Among the specialised historical atlases of recent years is the work dedicated to the development of the Slovak Evangelical Church (Kusendová 2010). After many years, the upcoming Academic Atlas of Czech History (AACH) is a comprehensive academic publication focused on Czech history from ancient times to the present.

2 Timeline of the work

The work of historians in the preparation of the content of the Academic Atlas of the Czech History begun in 2005. It was a part of the research project solved in the period 2005–2011 by the Institute of History, Academy of Sciences of the Czech Republic with topic on diversity of Czech historical space within a European context.

The conceptual work, and refinement of the structure and content of the atlas took place between 2005 and 2006, the team of authors was formed and the first maps with texts were drafted. In the following two years a manuscript of the atlas including maps and reconstruction drawings was prepared; in 2010 and 2011, the manuscript was reviewed by external critics and negotiations for financial support of the work started. The result of the work of professional historians consists of text, graphs, tables, and material collected for map preparation, usually in analogue form, or in simple graphic programmes.

A team of cartographers from the Department of Geomatics, Faculty of Civil Engineering was then gathered for digital cartographic processing of the maps before professional printing in the next phase of the project. The cartography for AACH began in 2012. As of June 2013, proof readings of all parts of this work are taking place and the atlas will be released in 2014.

3 Basic information about AACH

The title of the publication – the Academic Atlas of Czech History – was chosen due to the thematic focus of the work (Czech history), as well as the environment in which the authorial team worked on the scientific research project of the Academy of Sciences.

The goal of the atlas project is to process the results of historical research into Czech history in an international context. The atlas belongs to the small group of relatively few published scientific historical atlases. It reveals selected topics in modern Czech historical science after 1989 with connection to Europe, but especially Central Europe. The target user groups of the atlas are professionals in history, and the general public with an interest in history, historical geography, and related topics. The atlas arises as an extensive historical and cartographic work and it can also serve as a gift publication.

AACH presents the scientific activities of the Institute of History at the Academy of Sciences of the Czech Republic, however, it connects the results of historians with the output processed through creative collaboration with experts from several other institutions in the area of historical and social geography, demography, archaeology and other humanities and natural sciences. The conceptual background of the publication means that it can't present all the latest knowledge of historical sciences, studied in other research teams, universities, and other professionals. To be able to compare research issues from historical maps prepared by historians in the 1960s and again in the early 21st century, the atlas also indicates a list of maps in the Atlas of Czechoslovak History (1965).

The atlas is designed as a thematic atlas in the form of printed publications in A3 format (297 by 420mm) and it is formatted as an atlas encyclopaedia, containing text in individual sections supplied by maps, charts, pictures, photographs and other elements. The publication is written in Czech and contains a summary in English. The encyclopaedic concept of the work conforms to current trends (Voženílek 2008) and the upcoming work is significantly different in form to that of the historical atlas of 1965, where maps were strictly separated from the accompanying text. Editorial and publishing works are provided by the Academia Publishing House cooperating with the graphic studio. The atlas will be released as a printed publication; other (digital) forms are not intended.

4 The content of the atlas

The concept of the atlas content was designed by a group of authors from the Institute of History and is based on a conception of Czech history in an international, but especially Central European context. In addition to general research issues, selected several analytical probes of interesting historical processes displayed by reconstruction maps were selected for each period. In many cases, particularly in periods of modern history, the publication includes topics that were previously neglected or have been newly interpreted. The scope of the atlas was designed to be as balanced as possible, particularly with regard to the results of research projects of the Institute of History. The thematic content of the atlas is divided into the following sections:

- I. Prehistory (Stone Age – arrival of Slavs to the Czech lands): 27 pages, 8 chapters, 9 maps
- II. The Medieval Age (Samo Empire – reign of Jagiellons): 109 pages, 74 chapters, 82 maps
- III. The Early Modern Period (Habsburg dynasty – the end of the 18th century): 103 pages, 50 chapters, 86 maps
- IV. Top Modern Age (Napoleonic War – the disintegration of Austria-Hungary): 168 pages, 90 chapters, 111 maps
- V. Modern History (establishment of Czechoslovakia — the present): 88 pages, 40 chapters, 77 maps

The numbers above represent the maps created for the atlas using modern methods of digital cartography. The atlas also contains a number of thumbnails or reproductions (images) of previously issued maps relating to the topics of chapters. In addition to the thematic sections that present the main content area of the atlas, the publication includes an introductory section, geographical gazetteers and name indices at the end as well (see Fig. 11.1). Related references are included in each of the subchapters.

5 Cartographic process of making AACH

5.1 Organisational issues

For each of main thematic sections of the atlas a scientific editor–historian was appointed; the number of authors (of both text and graphics) for each chapter is considerable (experts from the Institute of History and from other institutions). The management team of historians was formed under the leadership of the Director of the Institute, Eva Semotanová. The editors of the sections were appointed: Martin Gojda (Prehistory), Robert Šimůnek and Josef Žemlička (The Medieval Age), Eva Chodějovská and Jiří Mikulec (The Early Modern Period), Aleš Vyskočil and Jan Hájek (Top Modern Age), Petr Prokš (Modern History).

Maps were prepared at the Department of Geomatics, Faculty of Civil Engineering by a team including Jiří Cajthaml (Prehistory and the Medieval Age; leading cartographer), Pavel Seemann (the Early Modern Period; design and preparation of the map key, cartographic revision of all maps), Tomáš Janata (Top Modern Age; design and preparation of specialised point symbols) and Růžena Zimová (Modern History). During the mapmaking, particularly vectorisation of the manuscript's background materials, other members of the department and several students made contributions.

Cooperation between cartographers and authors–historians during atlas creation is essential. In the case of AACH this cooperation was launched at an advanced stage of the project, when the manuscript of the atlas and most of manuscript maps and cartographic data had been prepared. Best practice in the creation of cartographic projects say that cartographers should work on formulating the concept of cartographic works and on the principles of map making (Pravda 2001, Slocum 2005, Voženílek 2011). This was not possible for this publication.

Due to relatively long intervals between manuscript creation and cartographic works the authors of sketch maps often had to modify or supplement the current state of knowledge in the field, or add or change maps. Due to the large number of authors maps were consulted mostly by editors of the individual sections from the Institute of History.

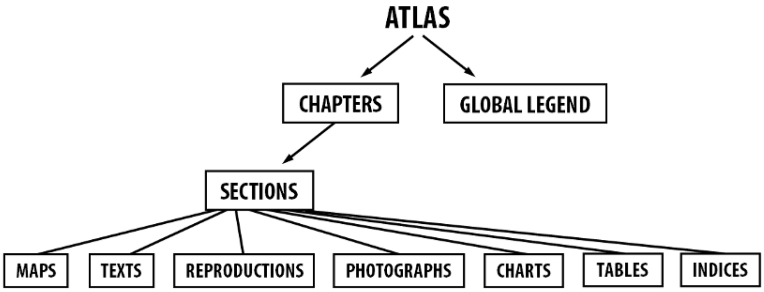


Fig. 11.1 Structure of the atlas

5.2 Materials for map creation

The data used in the processing of maps was diverse. In the authors elaborate concept maps the thematic content was often schematically hand-drawn onto a copy of a formerly published topographic or thematic map. In addition copies of previously published thematic maps were used; sometimes there was only a verbal description of the content of map elements (the list of thematic settlements or localities, line routes description of phenomena, etc.) or simply a reference to some existing map. Sometimes, digital thematic data (e.g. statistical data) was available. If the data was in digital form, it was usually in raster format; graphical data in vector form was available only in a few cases. A large portion of documents were delivered in analogue form and then scanned in a standard density (300 dpi).

5.3 Scale of maps

Atlas maps show the territory of the Czech Republic in different historical periods, or present thematic phenomena related to a larger, smaller or completely different area. In many cases it was therefore not possible to set a single series of scales due to the spatial and content diversity of the maps. The scale often had to be selected individually. Most often, maps at a scale of 1 : 2 million or similar were created; this corresponds approximately to a display of the Czech Republic in a map frame of width 265 mm. The structure of map scales is approximately as follows (using the scale number M, as a percentage of the total number of about 370 maps):

- M up to 200,000 ... 14%
- M ranging 200,000 – 1 million ... 6%
- M ranging from 1 to 2 million ... 51%
- M ranging from 2 to 4 million ... 16%

- M greater than 4 million … 13%

5.4 Size of maps

Due to differences in the range of the displayed areas (Europe, Central Europe, the former Czechoslovakia, the Czech Republic or smaller historical sites) and various requirements of the authors for detail in the subject represented, in many cases it was not possible to determine the dimensions of the map in advance, nor to introduce a simple variety of map frame sizes. The scope of the work process and the diverse character and quality of materials supplied did not allow the creation of a template for the final layout of atlas pages as recommended by the principles of atlas cartography (Voženílek 2011).

The maximum width of the map frame (265 mm) was determined according to the publication format (A3) and the design of the graphic concept of the atlas prepared by a professional graphic studio and recommendations to use one of several proposed sizes of map field were formulated as far as possible. Approximately half the total map frames have a dimension of 265 by 177mm, roughly a fifth of maps have a square frame, 177 by 177mm, and there are other larger or smaller sizes of map frame. The dimensions of the maps were chosen according to their spatial extent, appropriate scale for the plot and authors' themes as required, taking into account the general composition of the atlas pages.

5.5 Cartographic projection

As default, the Albers conic equivalent projection was chosen. The choice of the base meridian and undistorted parallels took place separately for each map and depended on the extent of the displayed area. For the frequently occurring extent of the Czech Republic at a scale of 1 : 2 million the parallels 45° and 55° N along the central meridian 16° E were chosen. Several maps (especially larger scales were made in other cartographic projections, mainly in view of the original data. The creators anticipated any implementation of cartometric or other geometric analysis of the content of the atlas maps, therefore it did not seem crucial to minimise the distortion of the maps caused by cartographic projection.

5.6 Map key

Creating a map key for such a large and diverse set of thematic maps presented a special issue. A map should fulfill its cognitive function (Pravda 1997). Various aspects of geovisualisation and map design are discussed regularly at scientific

events and/or in publications (e.g. Dykes et al. 2005, Kraak and Ormeling 2010). Nevertheless, the heterogeneity of historical topics involved in this historical atlas, the complexity and details of authors map concepts (and a demanding time schedule) made the process of map element symbolisation relatively difficult.
At the beginning of the cartographic work, a set of symbols for elements found in most maps was designed by the team of cartographers. These were particularly elements of topographical background, but also of common thematic features: areal (a scale for colour hypsometry, water bodies, colours for thematic areal elements such as areas of states, the colour scale for cartograms or diagram maps), line (various types of borders (Seemann 2012), hydrography, geographical grid, communications, thematic line elements such as those for military campaigns, front lines, etc.) or figural (settlements, castles, churches, monasteries etc.). Other symbols were gradually complemented by processing the respective map, the content of which was sometimes very specialised. The choice of colours was based on general cartographic rules (Brewer 2005, Voženílek 2011); the colours and parameters of figural signs were discussed and agreed with the graphic studio that designed and also provided fonts for map labels and all the text in the publication.

5.7 The composition and preparation for printing

The graphic and typographic solution, including the composition of the atlas pages, was designed and compiled by the graphic studio. Each map is surrounded by a thin map frame (line width 0.75 pt); thematic map legend and a simple graphic scale bar was placed the layout by the graphic designers within, usually out of the map frame (it was not always possible to put the legend on the same page of the atlas). A general list of symbols, containing characters mainly used to render the topographic base of the maps, is found in the opening section of the atlas. The name of each map was set by the authors–historians; sometimes it was necessary to refine or modify it. Numerical labelling of the maps corresponds to their inclusion in thematic sections. Author names (for text chapters and maps) are shown at the beginning of each thematic subchapter or through a refinement with the respective map. Supplementary compositional elements contain the north arrow used in large-scale maps without a geographical grid. Print layers of maps were exported as PDF (1200 dpi), or alternatively EPS (300 dpi) files for maps with colour hypsometry and hill shading, in the CMYK colour space. During the whole process, special attention was paid to structured, systematic and regular backup of all data. Some examples of maps in the atlas colours are available in the publications by Cajthaml (2013), Seemann (2013) or Seemann et al. (2013).

5.8 Some aspects of processing the maps

ArcGIS Desktop 10.0 software was chosen as the basic tool for digital map processing. It offers a number of cartographic tools for creating high-quality maps. Its advantage is the potential for a database-oriented store of map layers. An important issue was to use the appropriate attributes in order to facilitate correct map content depicting various time periods (e.g. the year of the water dam construction). Cartographic visualisation was performed with the standard methods of ArcGIS, without use of cartographic representations as that would have required more time which unfortunately was not available. In several cases, the cartographic program OCAD 10 was used, and final processing was performed using Adobe Illustrator.
The compilation of the topographic background of maps was mainly based on freely available reference data layers (SRTM30, CleanTOPO, ETOPO, Natural Earth, EUROSTAT), which were revised, supplemented and modified. For each specific map and its topographic component, the layers of thematic content were created, mainly by vectorisation of raster background. The creation of a data model in ArcGIS consisted of the design of feature classes within the dataset of each map, while generally usable reference data (e.g. altimetry, geographical grid, hydrography, administrative boundaries) was stored in a separate reference dataset (see Fig. 11.2) in two versions with different degrees of generalisation, usable in less detailed resolution for Europe or at larger scales for Central Europe or territory of the Czech Republic. The data model enables tracing of individual map data and sharing the layers between different maps. It represents a valuable database, usable both for supplementing further content of the maps and for creating any other similarly oriented cartographic work.

Cartographic visualisation of spatial phenomena was in some cases considerably complicated. It was often necessary to display various kinds of border lines, identification and movement lines, which were furthermore related to different time stages of the historical events presented. In this respect, the most complicated were the maps showing combat events in which it was necessary to distinguish the lines/arrows symbolising the movement of combat troops with regard to a variety of attributes, such as membership of the parties of the conflict, type of movement of combat troops (attack, advance, transfer, retreat, leaving), its intensity (major, minor, unresolved attack), military hierarchy (army, brigade, division, group, etc.), type of weapons (aircraft, tanks, etc.), time period of the war. Quantitative and qualitative attributes of these phenomena have been expressed by diversely selected parameters of linear features (structure, thickness, colour, orientation) in combination with the location of various end-line symbols (arrowhead single or double), adding another point symbol within line element or additional text or abbreviation (see Fig. 11.3).

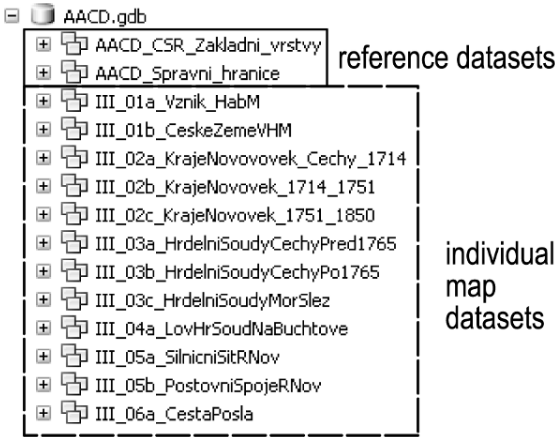


Fig. 11.2 Structure of the geodatabase

Delimitation of areal features in the map was depicted by coloured areas corresponding to the map key; for overlapping areas hatching was used. For the areas of countries, colour hypsometry in combination with hill shading (for maps containing figural and linear thematic elements) or solid colour fill with hill shading (for maps containing areal thematic elements) was applied. In maps of larger scales (approximately 1 : 500,000 and more detail), the relief is not included at all, or only in the form of more significant spot heights.

Czech geographic names were used for labelling geographic features; the official name in the language of the country was assigned to the settlements, if necessary supplemented with the name belonging to the period of historical event, such as Zlín (Gottwaldov), Chemnitz (Karl-Marx-Stadt). The use of exonyms represents a methodically serious issue that had to be solved at both a general level and when taking into account each specific map. With regard to the cultural–historical importance of exonyms, the names of settlements related to the topic of the map are accompanied by the Czech doublets in brackets, e.g. Dresden (Drážďany),Wrocław (Vratislav), Graz (Štýrský Hradec), and exceptionally also by doublets in a foreign language, such as Oświecim (Auschwitz, Osvětim) (see Fig. 11.4). The UN database of geographic names (UNGEGN 2013) and the List of Czech exonyms (Beránek et al. 2006) were used as sources of valid data. The cartographers, in collaboration with historians, tried to adapt the contents and labelling of maps in order to avoid overfilling the map image.

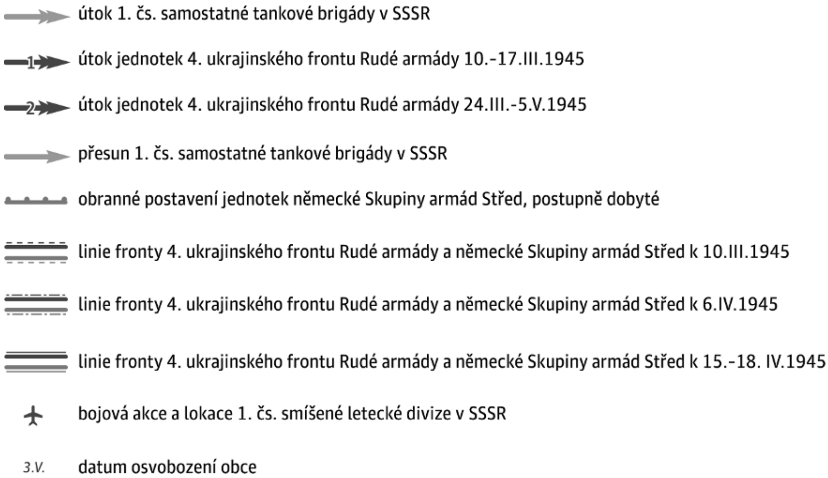


Fig. 11.3 Example of map symbols for combat events

6 Conclusion

The Academic Atlas of Czech History constitutes a unique multidisciplinary publication, which presents the results of the research of historians of the Academy of Sciences processed under cooperation with experts from several other departments in the areas of historical and social geography, demography, archaeology and other humanities and natural sciences.

In the process of the creation thematic historical maps, a close synergy of cartographers with the authors of professional thematic content is really essential. For quality processing of maps, it is important to understand the authors (historians) and to treat the historical phenomena correctly. Cartographers need not only the appropriate knowledge of and experience in map creation, but also some degree of knowledge about history, commonly used expressions and historical context. Experts in history (or other professional fields) are authors of the map content and therefore they are responsible for the factual accuracy of cartographic presentation of topics. Their effort to include correct, detailed and comprehensive historical content in the maps was sometimes found to be in conflict with what it was really possible to depict in a map of defined size and scale while respecting the principles of cartographic expression.

The processing of maps in ArcGIS software was carried out in the standard way but was not without several problems. It is worth remembering that the software does not always render graphical elements absolutely correctly (exported graphics are slightly different). It was also necessary to process the legends of size scales for diagram maps manually, as ArcGIS does not provide the tools to do so. The exported layers of maps also have unexpected outcomes in some settings and therefore everything should be carefully examined. In general, this software could be certainly recommended for atlas creation. With compliance to the basic principles, the job is quick and efficient. It is very important to coordinate the appropriate procedures with a graphic studio providing professional printing of the publication.

Fig. 11.4 Example of map with exonyms

Based on the experience of the cartographic processing of maps for the AACH, it is necessary to draw attention to two important aspects: firstly, the need to involve cartographers in the preparation of the atlas at the very beginning of the project, and secondly, the emphasis on quality and unity of base sources for creation of maps. The concept of the atlas should be formed by authors-historians, always in cooperation with cartographers. In the initial phase, the rules used for the preparation of data for map creation should be mutually agreed. In the case of AACH, there was a friendly cooperation between authors and the team of cartographers at an excellent level. Nevertheless, the involvement of cartographers in the project of the atlas from the beginning could have solved many problems and positively influenced the time required for map processing. Uniform data processing for cartographic production would require considerable efforts before starting work on authorial originals. This practical problem particularly affected the duration of the preparation of the atlas and caused uncertainty in its processing, which is usually caused by the financial resources for its development. For the team of cartographers, the work of processing the Academic Atlas of the Czech History was a very valuable experience which it will be possible to further develop in the future.

References

Beránek T, Boháč P, Šrámek R, Drápela MV, Harvalík M, Liščák V, Šimůnek R (2006) Index český ch exonym. Standardizované podoby, varianty. (List of Czech Exonyms Standardized Forms and Variants). Praha: Český úřad zeměměřický a katastrální, 2006

Brewer C (2005) Designing Better Maps: A Guide for GIS Users. Redlands: ESRI Press, 2005

Cajthaml J, Janata T, Seemann P, Zimová R (2013) Designing Maps for a New Thematic Atlas of the Czech History. In: Proceedings of the 26th International Cartographic Conference, Dresden, Germany. International Cartographic Association, 2013

Dykes J, MacEachren AM, Kraak MJ (2005) Exploring Geo-visualization. Amsterdam: Elsevier, 2005

Kraak MJ, Ormeling FJ (2010) Cartography: Visualization of Geospatial Data. Harlow, Pearson Education, 2010

Kusendová D (2010) Historical Atlas of the Evangelic Church in Slovakia – information (in Slovak). In: Aktivity v kartografii 2010. Bratislava: Kartografická spoločnosť Slovenskej republiky, 2010

Pravda J (1997) Language Aspect of Map Representation. In: Proceedings of 18th ICA International Cartographic Conference. Stockholm: Swedish Cartographic Society, 1997

Pravda J (2001) On General Principles of Map Design (in Slovak). In: Geografický časopis, 53. Bratislava: Geografický ústav Slovenskej akadémie vied, 2001

Purš J et al (1965) Atlas československých dějin. (Atlas of the Czechoslovak History; in Czech). Praha: Historický ústav ČSAV, Ústrední správa geodézie a kartografie, 1965

Seemann P (2012) Cartographic Representation of Border Lines in the Academic Atlas of the Czech History (in Czech). In: Sborník Digitální tech- nologie v geoinformatice, kartografii a dálkovém průzkumu Země. Praha: ČVUT – Fakulta stavební, Katedra mapování a kartografie, 2012

Seemann P (2013) Academic Atlas of the Czech History — Summary of the Contents and Experiences from the Production (in Czech). In: Sborník Digitální technologie v geoinformatice, kartografii a dálkovém průzkumu Země. Praha: ČVUT – Fakulta stavební, Katedra geomatiky, 2013

Seemann P, Cajthaml J, Janata T, Zimová R (2013) Czech Lands within the Habsburg Monarchy during 17th and 18th Centuries. In: ESRI Map Book, Volume 28, pp. 60–61, Redlands: ESRI Press, 2013

Semotanová E (2003) Three Historical Maps of Aleš Pařízek from the End of 18th Century (in Czech). In: Historická geografie, Volume 32, pp. 69–92 Praha: Historický ústav AV ČR, 2003

Slocum AT, McMaster RB, Kessler FC, Howard HH (2005) Thematic Cartography and Geographic Visualisation. Upper Saddle River: Pearson Prentice Hall, 2005

UNGEGN World Geographical Names (2013) http://unstats.un.org/unsd/geoinfo/geonames/. Accessed 30 October 2013

Voženílek V (2008) Atlas of Climate in the Czech Republic (in Czech). In: Kartografické listy, Volume 16, Bratislava: Kartografická spoločnosť Slovenskej republiky, 2008

Voženílek V et al (2011) Methods of Thematic Cartography — Visualization of Spatial Phenomena (in Czech). Olomouc: Univerzita Palackého v Olomouci, 2011

Part III
Context-Oriented Cartographic Visualisation

Cognitive Aspects of Geovisualisation: A Case Study of Active and Passive Exploration in a Virtual Environment

Asli Yilmaz, Serkan Kemec, H. SebnemDuzgun and Murat P. Cakir[1]

Abstract Visualisation of geospatial information is a key issue as it is a bridge between rich and high level spatial information, and the users of this information that helps to support decision making, management and operations. In recent years, new developments in technology have provided new methods and platforms that enable the innovative visualisation of geospatial information. Disciplines such as information visualisation, scientific visualisation, human-computer interaction and cartography have been integrated by researchers in order to generate geovisualisation strategies. Many aspects of Geographic Information Systems (GIS) could be improved by greater attention to the research on cognitive science. Operators, users and decision makers using GIS deal with large and complex geospatial data, and so the way geospatial information is visualised affects their perception. This explorative study suggests a systematic methodology for testing and analysing perceptions of geospatial information, considering different types of explorations and visualisation scenarios in a 3D virtual city environment. After making behavioural and fNIRS analyses, the study discusses how cognitive issues may differ in these scenarios.

1 Introduction

Geographic Information Systems (GIS) allow large amounts of data to be efficiently stored, manipulated, analysed, and displayed. They provide a systematic combination of geographic information to support their users' decision making processes, management and operations. Spatial decision support is one of the important areas assisted by GIS.

A. Yilmaz, H. Sebnem Duzgun, M. P. Cakir
Middle East Technical University,
Ankara, Turkey
e-mail: aslimail2me@gmail.com

S. Kemec
Yuzuncu Yil University,
Turkey

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_13,

"It means computerized assistance to people in the development, evaluation and selection of proper policies, plans, scenarios, projects or inventions where the problems have a geographic or spatial component. This refers to both long-term decision making such as planning for sustainable places, mitigating hazards, infrastructure management, strategic business planning and short-term critical decisions such as emergency response and resource logistics" (Andrienko et al. 2007).

The visualisation of geospatial information is important for effective decision making (Zlatanova et al. 2002, Kemec and Duzgun 2006, Zlatanova 2008). Thus, the visualisation of geospatial information, termed geovisualisation, is an emerging tool for spatial-decision support which creates synergy between computational techniques and human capabilities (Andrienko et al. 2007).

MacEachren and Kraak (2000) present the results of an international collaboration to set a comprehensive research agenda which expressed research challenges as geo visualisation. Four themes appeared: the representation of geospatial information, integration of visual with computational methods of knowledge construction, interface design for geovisualisation environments, and the cognitive and usability aspects of geovisualisation. As stated in the fourth theme, involving cognitive and usability issues, it is important to understand how people react to fully three dimensionally enhanced environments, how they deal with complex and large amounts of information and how effective are the navigation tools that provide the information. This theme emphasised that in order to understand the factors that determine the success or failure of geovisualisation, the perspectives of cognitive and usability issues associated with geovisualisation should be extended and new metaphors, theories, tools and methods should be generated (Mac Eachren and Kraak 2000).

In recent years, new developments in technology have provided new methods and platforms to enable the innovative visualisation of geospatial information. Among these, 3D geospatial environments have been employed in an increasing number of applications, including city planning, city marketing, tourism and facility management (Altmaier and Kolbe 2003). MacEachren and Kraak (1999) suggested that 3D geospatial environments are super environments, as they provide user experiences that are not only visible but also invisible. The use of 3D geospatial environments is categorised into three groups by Bleisch (2012): 3D representation of abstract data, 3D representation of the real world and 3D representation of the real world and abstract data in combination. In the meantime, virtual 3D city models are rapidly increasing in number with explicit semantics, topology, and thematic information (Döllner 2009). They have become essential computational tools as they allow 3D analysis, simulation, navigation, communication and management (Döllner 2009). Examples of the use of 3D city models include city walk-throughs or fly-throughs to show what a new building would look like in situ, whether light or a view will be blocked by a new structure (Ellul and Joubran 2012). According to Petzold and Matthias (2011), 3D city models are generally more useful if they include additional data which can be analysed as a 3D representation of the real world.

Integrating cognitive evaluation of geovisualisation in a 3D virtual environment is the focus of this study. Cognitive aspects of conveying geospatial information cover human perception, memory, reasoning, problem-solving, communication and visualisation (Montello 2005). Researchers address many issues involving both geospatial information and cognition, such as "how geographic information is learned and how this learning varies as a function of the medium through which it occurs (direct experience, maps, descriptions, virtual systems, etc.)", "What are more natural and effective ways of designing interfaces for geographic information systems?" "How do people develop concepts and reason about geographical space, and how does this vary as a function of training and experience?", "how complex geographical information can be depicted to promote comprehension and effective decision-making, whether through maps, models, graphs, or animations?" etc. (Montello 2005).

Geovisualisation cannot be properly undertaken without understanding the human perception of space. Which types or ways of visualising geospatial information create the minimum cognitive effort is a critical question. The traditional 2D use of spatial information is becoming out of date. 3D visualisations and the use of Virtual Environment (VE) technology to assess spatial navigation have become increasingly common. According to Kolbe et al. (2005), 3D graphical representations significantly reduce the amount of cognitive effort required, and improve the efficiency of the decision-making process. Virtual reality techniques also provide better perception and comprehension of complex 3D structures (Beck 2003). "Meanwhile, there is a further trend towards 3D virtual reality GIS. These systems can represent and handle complex three-dimensional objects like buildings and allow for real-time visualization applications. The rapid developments in the field of computer graphics have also supported the use of virtual reality components within standard GIS" (Haala 2005).

This chapter presents an explorative study, and suggests a systematic methodology for testing and analysing the perception of geospatial information, considering different types of explorations and visualisation scenarios in a 3D virtual city environment. After making behavioural and fNIRS analyses, the study discusses how cognitive issues may differ in these scenarios.

2 Methodology

According to Montello (2005), one of the priority areas for research into the cognition of GIS in the next 3-5 years is the exploration of GIS in the possible varieties of VE technologies, which is the main focus of this study. This study investigates how different methods of geovisualisation in a virtual environment affect human perception. By comparing participant explorations through active-passive navigations and the use of static-dynamic graphic representations of spatial information, cognitive effort results are discussed. In this study, a 3D model of an urban environment with mixed building types and height characteristics was created in a virtual environment for non-expert GIS users, and its use in a spatial decision support system for earthquake risk in an urban area was presented and analysed. The 3D city model was used to visualise each building's earthquake risk levels, information about the characteristics of the buildings and their spatial positions.

2.1 Participants

The participants of this study were five non-expert users of GIS. They can be characterised according to age, gender, profession, computer experience and navigation experience of virtual environments (Table 13.1). Three were female; their mean age was 23.2; the focus of their university study was related to computer science; their mean computer experience was 12.2 years; and they had all taken part in navigation in a virtual environment.

Table 13.1 Participant characteristics

	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5
Age	23	23	22	28	20
Gender	Female	Female	Male	Female	Male
Profession	Computer Education and Instructional Technology	Computer Science	Computer Engineering	Computer Education and Instructional Technology	Computer Engineering
Computer Experience	11 years	12 years	13 years	15 years	10 years
Virtual Env. Navigation Experience	Yes	Yes	Yes	Yes	Yes

2.2 Virtual Environment

Unity Version 3.0 software was used as the tool to create the virtual environment of the study. Unity is a game development tool with rich interactive 3D content allowing 3D objects to be built into and animated in a scene. 3D meshes from other sources that can also be imported into the scene that is created. With its open source code property, unique scenarios can be created and actors can be brought to navigate the built environment. The actor's size, type, speed etc. can be changed, and they can gain abilities such as walking, running and flying. The project can be exported to be played in a Windows standalone mode.

In this study, a 3D city model of part of the Eskisehir metropolitan area was imported into the Unity scene and a satellite image of the area was projected onto the terrain (Figure 13.1). Four separate tasks were created in the same project and virtual environment. Four Windows with standalone modes for the tasks were exported, in which different screen resolution and graphic quality options could be selected.

Fig. 13.1 Screenshot from the scene

An actor whose speed was 200m/sec for each task was put into the scene to perform the navigation. For each task, it was expected that a target building would be found and information about the building be gathered. This information would include s*patial* and *tabular* data for the building as well as information on the pop-up window which appeared at the bottom right of the screen once the building was clicked. Information about the pop-up window included textual information about the characteristics of the building and earthquake risk level information (Figure 13.2). Once the target building was clicked and the pop-up window appeared, there was a 20 second time limit before the entire screen turned black.

Navigation was achieved using arrow or W, A, S, D buttons. Zoom in and out was enabled using a mouse wheel. In order to understand where the target building is, a top view of the scene is basically shown in a window on at the bottom right of the screen left-down. The building is also highlighted from time to time (Figure 13.2).

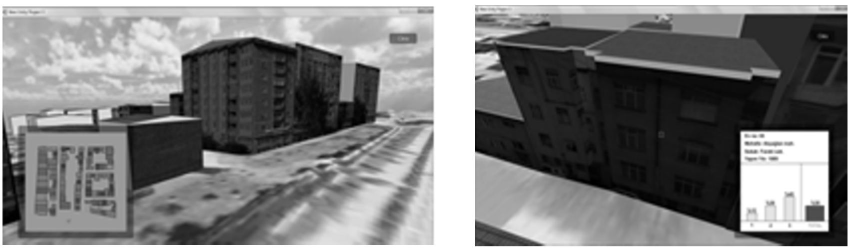


Fig. 13.2 Highlighting and pop-up window

It is important to indicate that in the first task, a passive exploration of the city was planned, where the unity system automatically navigated through the city and found the building. In the following tasks, exploration was active; navigation tools were used. In the first three tasks, graphics for the pop-up window were static, but for the last task, it was designed as a growing chart animation. Three combinations of exploration were thus created: "passive-static" (Task 1), "active-static" (Task 2 and 3) and "active-dynamic" (Task 4). "Active" tasks were formulated so as to understand the behaviours of participants in space. One passive task (passive-static) was created to understand whether the participants retrieved information changes when they did not actively navigate through the virtual environment. This is analogous to being a person driving a car, or being a person sitting next to the driver, who can better retrieve information about the route. This was explored through a comparison of active-static and passive-static tasks. Another comparison explored active-static and active-dynamic tasks so as to understand whether graphics as animated or as static in a pop-up window affects the better perception and retrieval of information.

Two active-static explorations were used to allow the creation of similar building and position characteristics in Task 1 and Task 4, for better comparisons of active-passive and static-dynamic explorations. The first two tasks' target building and position characteristics were selected to compare and similar logic was also applied for the last two tasks. What is meant by building characteristics was building type, number of floors and material. What is intended by position characteristics was whether the building was located in the main street or side street.

2.3 Data Collection Procedure

Each participant was welcomed and a short training session described what would happen during the test session. The training sessions allowed the participants to get used to the scene environment and navigation tools. Test sessions were held at the Middle East Technical University (METU) Human Computer Interaction Laboratory and the MODSIMMER Human Factors Laboratory. The participants were observed without being disturbed. Free active window recorder software, Webinaria, was used to record the participants' screen movements. The participants were informed of this recording at the beginning of the test and their consent was obtained.

Functional optical brain imaging technology, Functional Near-Infrared Spectroscopy (fNIRS), was used during the test sessions. It is a safe and portable system and offers real time monitoring of brain activity. Studies of fNIRS over the last decade indicate that it "provides veridical measure of oxygenation and blood flow". Recent findings suggest that "fNIRS can effectively monitor cognitive tasks such as attention, working memory, target categorization, and problem solving" (Izzetoglu et al. 2005).

fNIRS measures changes in the concentration of oxygenation and deoxygenation mainly in the front of the brain in the dorsolateral prefrontal cortex; more precisely the following Brodmann's areas; BA9, BA10, BA46, BA45, BA47 and BA44 (Izzetoglu et al. 2005; Merzagora and Izzetoglu 2008). A fNIRS sensor was attached to the participant's forehead and monitored, connected directly to a portable computing device that recorded the participant's data as they were doing each task. The way brain activity was being affected by certain tasks could be monitored.

After each task was finished, participants were asked to fill a questionnaire in which five questions were asked for each task. Four question sets (20 questions in total) were prepared for each task. For each task, the questions aimed to collect data about the memories of the participants as related to *spatial* (one question for each task) and *tabular* information (four questions for one task). The reason for this categorisation comes from the nature of GIS. In GIS, information is categorised as *tabular* and *spatial.* For instance, questions such as "Was there a building besides the target building? If there was, was it on the left side or right side?" or "Where was the building located?, Is it in the main street or side street? In the corner or on side?" were asked to understand the retrieval of information related to space. Questions such as "What was the colour of the building?" or "What was the year of construction of the building?" were asked to understand the retrieval of *tabular* information.

The questions were designed in a hierarchic manner, from general to detailed. For instance, the first question was a simple general question, such as asking whether the height of the building could be described as short, mid or high. The final question was more difficult and specific, such as asking for an assessment of risk level in numbers seen on the graphic. There was one additional question that that asked participants to describe any other details that they remembered. This data would also reveal their attention during tasks. The data was collected in order to understand the differences and similarities between tasks and participants through behavioural data analysis.

3 Data Analysis

3.1 Behavioural Data Analysis

In order to conduct behavioural data analysis, video records of the participants were analysed. Data regarding the time it took participants to complete the tasks, memory errors, and navigation paths was generated. By *memory errors,* the participants not recalling the right information (choosing the wrong answer or not answering the question) was meant. In the first task, completion time was equal because it was a passive exploration task generated by the Unity system, and participants only watched the exploration. It took 37 seconds, but other tasks took longer. Completion times for each participant can be seen in Table 13.2. Completion times for Task 3 and 4 were similar. It was easier for participants to find the 2nd target building than to find the other buildings. This may be because the building was bigger in size and located on the corner of the main street. Three of the participants spent longer on Task 4 than on the previous tasks. This cannot be connected to the animated graphics designed for this task as the time limit for watching the graphics was the same. The mean completion time results did not seem to be similar. It should be noted that participant numbers were too small for accurate statements of comparison.

Table 13.2 Time to complete the tasks

Participant	Task 1 (sec)	Task 2 (sec)	Task 3 (sec)	Task 4 (sec)	Mean (sec)
1	37	96	104	84	80.25
2	37	71	116	140	91
3	37	70	79	96	70.5
4	37	64	63	65	57.25
5	37	102	77	70	71.5
Task Mean (sec)	37	80.6	87.8	91	

The answers to each questionnaire were analysed in terms of memory errors related to *spatial* and *tabular* information (Table 13.3). In the table, memory error rate (error /number of questions) given. Memory errors related to *tabular* information were at least in Task 1. This might be because participants were able to pay attention to information more easily during passive exploration. On the other hand, most memory errors were related to *spatial* information in this task. This might be because participants could not actively control the navigation and understand *spatial* information about the target building. The results suggest that errors in retrieving *spatial* information were not related to the positions of the buildings because the second and fourth target buildings were positioned completely differently. One was at the corner of the main street, and one was on the cross street. According to these results, memory errors were also not related to the size or shape of the target buildings. Errors increased in parallel with the level of detail in the question. This type of errors may have been about the graphic value of risk levels, the name of the street, number of floors etc. In Task 4, the animation of the risk levels might have had bad effect on obtaining values because two participants made errors related to this, when they had not done so in previous tasks. This could be better understood if the participant size was higher.

Table 13.3 Memory errors

	Memory error ratio related with *tabular* information	Memory error ratio related with *spatial* information						
Participant	Task 1	Task 2	Task 3	Task 4	Task 1	Task 2	Task 3	Task 4
1	1/4	1/4	0	0	0	0	0	1/1
2	0	1/4	2/4	1/4	1/1	1/1	0	0
3	0	0	0	2/4	0	0	0	0
4	0	1/4	0	0	0	0	0	0
5	0	0	0	0	1/1	0	0	0

Paths travelled by the participants can also be analysed to follow their behaviour during navigation in the environment. As the navigation paths of the participants are seen from the top view, they tended to prefer the shortest paths to get to the target buildings (Figure 13.3). The starting point was the same for each task, as seen in Figure 13.3, and their end points were next to the target buildings. During Task 1, there was passive navigation. For Task 2, all participants chose the shortest path. Participant 1 walked a longer distance than the others because she passed the building while navigating and turned back. During Task 3, four of the participants followed the shortest path but Participant 2 did not. During the last task, three participants followed the shortest path but two did not. This might have been because the target building was placed close to the middle of the top view map. This path view indicated that it was harder to choose the shortest paths to the buildings that were on cross streets, where there were narrow spaces, rather than to buildings on the main street where the spaces were larger.

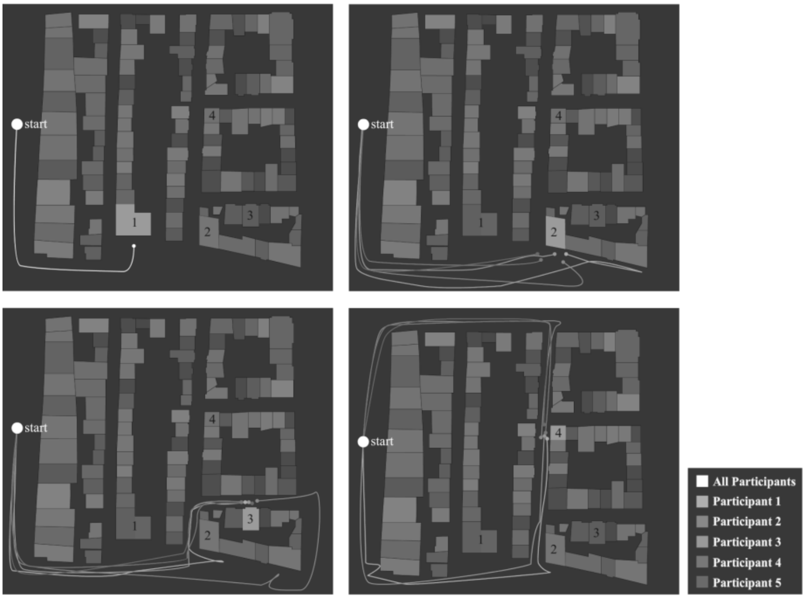


Fig. 13.3 Participant navigation paths (top view)

3.2 fNIRS Analysis

As noted above, there is a relationship between metabolic activity and oxygenated haemoglobin in the blood vessels in the prefrontal cortex. For this study, the oxygenation data collected by 16 channel sensors were filtered and prepared as mean oxygenation data. Increased oxygenation suggests that participants put more cognitive effort into performing the task. Since the study focuses on brain functions such as attention and memory for oriented spatial navigation tasks, the responses from fNIRS are supplementary in making better comparisons of tasks and task conditions.

The mean oxygenation data for all tasks can be seen in Table 13.4. Participant 1's mean oxygenation data is the highest, and otherwise the mean oxygenation data seems to be similar.

Table 13.4 Descriptive statistics/participant – fNIRS mean oxygenation

Participant	Mean oxy.	Std. Deviation
1	10.9201	.1798
2	28.1019	.5682
3	14.6812	.4413
4	12.4218	.8812
5	16.3323	.3705
Total	17.1925	6.5675

Oxygenation level increases with every following task (Table 13.5). Task 1, the passive exploration, generated, minimum oxygen levels in the participants. A sample t-test was generated to determine the mean difference in oxygenation between task groups performed by all participants. As the significance level is more than 0.05, there was no significant difference in the variation of the task groups (Table 13.6). In addition, left and right brain mean oxygenations and t-tests of these groups were also generated. No significant differences were found.

Table 13.5 Descriptive statistics / task –fNIRS mean oxygenation

Task	Mean oxy.	Std. Deviation
1	15.1262	5.9544
2	16.0453	5.7336
3	17.8781	7.0089
4	18.9667	6.7055
Total	17.1925	6.5675

Table 13.6 One sample t-test/task – fNIRS mean oxygenation

Task	t	df	Sig. (2-tailed)	Mean Difference	%95 Confidence Interval (Lower)	%95 Confidence Interval (Upper)
Task 1	5.084	4	.007	14.9299	6.7758	23.0839
Task 2	5.106	4	.007	14.5051	6.6177	22.3925
Task 3	4.784	4	.009	14.5094	6.0895	22.9293
Task 4	4.981	4	.008	14.3689	6.3591	22.3788

In order to understand differences between active and passive explorations and the use of static and dynamic graphic representations, two more t-tests were generated (Table 13.7 and 13.8). Tasks were grouped and task conditions were created. According to the results, with a significance level of more than 0.05, there was no significant difference between each task condition.

Table 13.7 One sample t-test/task condtion1 – fNIRS mean oxygenation

Task Condition	t	df	Sig. (2-tailed)	Mean Difference	% 95 Confidence Interval (Lower)	% 95 Confidence Interval (Upper)
Task Passive	5.577	4	.005	16.8782	8.4752	25.2812
Task Active	4.949	4	.008	14.5023	6.3666	22.6380

Table 13.8 One sample t- test / task condition 2 – fNIRS mean oxygenation

Task Condition	t	df	Sig. (2-tailed)	Mean Difference	%95 Confidence Interval (Lower)	%95 Confidence Interval (Upper)
Task Static Graphic	4.784	4	.009	14.5094	6.0895	22.9293
Task Dynamic Graphic	4.981	4	.008	14.3689	6.3591	22.3788

4 Results and Discussion

This explorative study suggested a methodological approach to testing and analysing different types of navigational experiences and different types of information visualisations in a 3D virtual environment. Participant behaviours were analysed by conducting two different explorations (passive-active) and two different information visualisations (static-dynamic). Using fNIRS, the differences between the metabolic activities of participants during the tasks were analysed.

Errors related to retrieving *spatial* information could be minimised by allowing participants to control the environment more actively. Dynamic graphics might have had a negative effect on perceiving the information, as participants might have paid attention to the motion itself rather than the information. The position of the buildings might have caused a number of memory errors in addition to affecting the navigation path preferences. On the other hand, size and type of the building did not create any difference in this study. Although there were minor differences that can be deducted from the behavioural data, there was no significant difference in cognitive load in terms of memory, planning or attention between the tasks and task conditions.

According to cognitive researchers, in terms of spatial navigation the hippocampus area, which is located in the medial temporal lobe of the brain, plays an important role. Analysis of the prefrontal cortex with fNIRS might therefore not be enough to determine the differences in activation. Also, especially some of the data from fNIRS voxels was eliminated as the headbands could not fit on participants' heads well, particularly for female participants, and caused 'dirty data'. However, fNIRS is a portable device and its adaptation to the test environment was simple. Its use in further studies is preferable, but analysis can be undertaken using eye-tracking methods.

To achieve statistically more robust results and to be ensured the assessments more participants are required. Comparisons between different types of users could be generated in further studies. The behaviour of expert GIS users may differ from that of the users who constituted this study (non-expert) as they more often operate with spatial content in a virtual environment. Sets of tasks including different geovisualisations could be designed and conducted. The development of new virtual environment tools, scenes, 3D objects, and the effects of texture, light and shading types should also be discussed in further studies.

References

Altmaier A, Kolbe TH (2003) Applications and Solutions for Interoperable 3D Geo-Visualization. In: D. Fritsch (ed.) Proceedings of the Photogrammetric Week, Stuttgart, Wichman Verlag 251-267

Andrienko G, Andrienko N, Jankowski P, Keim D, Kraak M, MacEachren A, Wrobel S (2007) Geovisual Analytics for Spatial Decision Support: Setting the Research Agenda. International Journal of Geographical Information Science 21(8), 839-857

Beck M (2003) Real-Time Visualization of Big 3D City Models, International Archives of the Photogrammetry. Remote Sensing and Spatial Information Sciences, 34 (5/W10)

Bleisch S (2012) 3D Geovisualization- Definition and Structures for the Assessment of Usefulness. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume I-2, 2012 XXII ISPRS Congress, 25 August - 01 September 2012, Melbourne, Australia

Döllner J (2009) Virtual 3D City Models as Computation Tools. Geoviz Workshop Dealing With The Contribution Of Geovisualisation To The Concept Of The Digital City, ICA Commission on Geovisualisation, , 3-5 March 2009, Hamburg

Ellul C, Joubran JA (2012) Preliminary Investigations into the Potential of Improving Rendering Performance of 3D Datasets using 2D Generalization. European Cost Action TU801, Final Conference

Haala N (2005) Towards Virtual Reality GIS. In Dieter Fritsch, (ed.) Photogrammetric Week'05, p.285-294, WichmannVerlag, Heidelberg

Izzetoglu M, Izzetoglu K, Bunce S, Ayaz H, Devaraj A, Onaral B, Pourrezaei K (2005) Functional Near-Infrared Neuroimaging. IEEE Transactions on Neural Systems and Rehabilitation Engineering,. 13(2) p.153-159

Kemec S, Duzgun S (2006) 3D Visualisation for Urban Earthquake Risk. In "Geohazards", Farrokh Nadim, Rudolf Pöttler, Herbert Einstein, Herbert Klapperich, and Steven Kramer Eds., ECI Symposium Series, Volume P7 (2006)

Kolbe TH, Gröger G, Plümer L (2005) CityGML - Interoperable Access to 3D City Models. Oosterom P, Zlatanova S, Fendel E. (eds.) Proceedings of the International Symposium on Geo-information for Disaster Management, 21-23 March 2005, Delft, Springer Verlag

MacEachren AM, Kraak M (1999) Virtual Environments for Geographic Visualization: Potential and Challenges. In Proceedings of ACM Workshop on New Paradigms in Information Visualization and Manipulation, ACM, Kansas City, p.35-40

MacEachren AM, Kraak M (2000) Research Challenges in Geovisualisation. Cartography and Geographic Information Science, 28(1)

Merzagora A, Izzetoglu K (2008) Optical Brain Imaging FAQ, http://www.biopac.com/FAQ-Details.asp?ID=464. Accessed 24 Nov 2013

Montello DR (2005) Cognition of Geographic Information. McMasteri RB & Usery EL (eds.), A Research Agenda for Geographic Information Science. p. 61-91. Boca Raton, FL: CRC Press

Petzold B, Matthias E (2011) Killerapplikationgeseucht.gis.Business,7, p.40-42

Zlatanova S, Rahman AA, Pilouk M (2002) 3D GIS: Current Status and Perspectives. In the Proceedings of the Joint Conference on Geo-spatial Theory, Processing and Applications, 8-12 July, Ottawa, Canada, 6p. CDROM

Zlatanova S (2008) SII for Emergency Response: the 3D Challenges. In: ChenJ, JiangJ, Nayak S (Eds.); Proceedings of the XXI ISPRS Congress, Part B4-TYC IV, Beijing, pp. 1631-1637

Assessing the Completeness and Positional Accuracy of OpenStreetMap in China

Shudan Zheng and Jianghua Zheng

Abstract OpenStreetMap (OSM) has been successfully applied all around the world, especially in developed western countries, but this is the first study of the quality of OSM data in China. Two data quality elements, completeness and positional accuracy, were chosen to conduct the assessment via a comparison against Baidu datasets. This chapter quantitatively depicts some characteristics of the distribution of OSM data based on the density of line and point features. The analysis showed that 71% of the OSM data was less detailed than the Baidu datasets, but on average 66% of OSM data was accurate. The OSM data for Beijing and Shanghai is most complete with high positional accuracy. Overall coverage was extremely poor, as more than 94% of the country consisted of 'incomplete regions' (regions with few or no data). However, OSM data has grown quickly, according to a comparison of three years, 2011 to 2013. More interestingly, OSM contained more detailed information in some poor areas, which could be an improvement over datasets provided for normal users by commercial or governmental agencies.

1 Introduction

Nowadays, geographical data is available to the public through increasingly dynamic geographical websites. As the Geoweb becomes more and more popular, a new movement, the crowed-sourcing movement, is developing in a similar way. Crowd-sourcing greatly reduces the time taken to obtain and disseminate geographical information. Among these 'Volunteered Geographic Information' (VGI) projects (Goodchild 2007), OpenStreetMap (OSM) which is a version of crowd-sourcing may be the most representative. OSM is a collaborative project, similar to Wikipedia, which was started in England in 2004 by Steve Coast (Ramm et al. 2011). It aims to create a fully free and openly accessible map of the world. Volunteers can create maps of their local areas, using local knowledge, more effectively than a mapping expert from a distant government agency. With its ad-

S.D. Zheng • J.H. Zheng (✉)
Xinjiang Education Ministry Key Laboratory of City Intellectualizing and Environment Modelling, Xinjiang University,
Urumqi, China

This work has been supported in part by NSFC of China (Project No. 41361084)
Jianghua Zheng: E-mail: itslbs@126.com Telephone: +86 13579880590
Shudan Zheng: E-mail: shudan.zheng@gmail.com Telephone: +46 2712713

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_14,

vantages of being open and free, the coverage, density and number of OSM users is quickly increasing worldwide (Figure 14.1[1]).

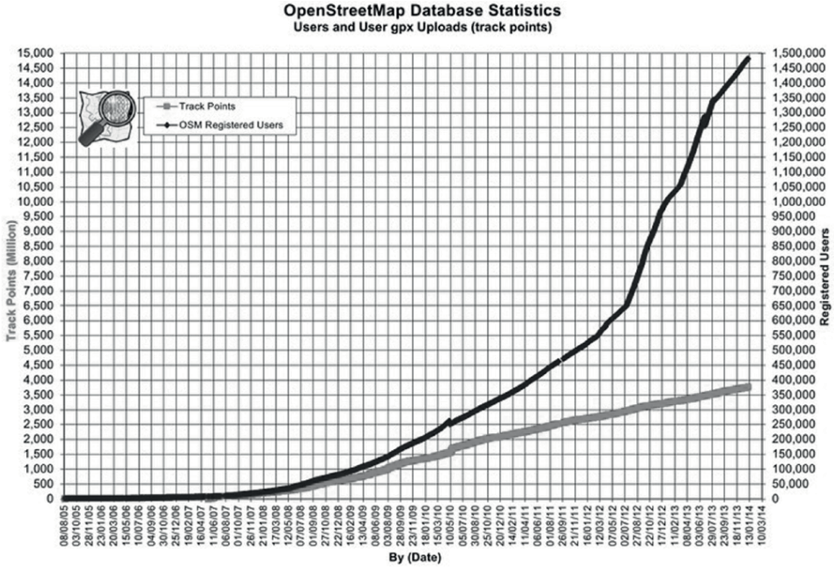


Fig. 14.1 Graph of registered users and growth in contributions to OSM on a monthly basis. OpenStreetMap had nearly 1,500,000 registered users in January, 2014, and the increase in data contribution continues to rise quickly.

There is no doubt that OSM is a very popular project as a means of collecting and providing timely and detailed geographic information for free. So far, it has gradually become a major database, and been successfully applied in different fields by many scientists in western counties, in projects such as 3D City Models (Over et al. 2010), pedestrian navigation (Jacob et al. 2009a, 2009b, Zheng et al. 2010), robot navigation (Hentschel and Wagner 2010), spatial queries (Yin and Carswell 2011; Codescu et al. 2011), and disaster logistics (Neis et al. 2010). Unlike the traditional authoritative geographic information which it potentially augments and even replaces in some cases, however, it carries no assurance of quality (Goodchild and Li 2012). Understanding and evaluating the quality of OSM data provided by volunteers before applying it to specific purposes or research is therefore a significant problem today.

Several studies have been made evaluating the quality of OpenStreetMap data against conventional geographic information sources. Haklay (2010a), for example, compared the quality of OpenStreetMap with data from the Ordnance Survey Meridian 2 dataset from Great Britain's national mapping agency. He analysed the positional accuracy, completeness and number of users per area of both OSM and OS Meridian data. He found that there was approximately 80% overlap in motorway objects between the two datasets. Importantly, OSM had captured approxi-

[1] OpenStreetMap. Http://wiki.openstreetmap.org/wiki/Stats. Accessed 9 Feb 2014.

mately 29% of the area of England over the past four years. Girres and Touya (2010) extended the works of Haklay to demonstrate that a larger set of spatial data quality elements (point, linear and polygon primitives for geometric accuracy, attribute accuracy, semantic accuracy, completeness, logical consistency, temporal accuracy, lineage and usage) and different methods of quality control had been used to assess the quality of French OpenStreetMap data against data captured from France national mapping agency (Instítut Géographique National, IGN). They found that responsiveness and flexibility were advantages of the OSM data, but that there were also some obvious shortfalls in heterogeneity and limitations in the possible applications. Zielstra and Zipf (2010) made a contribution to comparing OpenStreetMap and TeleAtlas data in Germany and found that the total length of roads in OSM was still smaller than in TeleAtlas. The growth rate of the entire street network in OSM is tremendous, however, as the differences between two datasets reduced from 29% to 7% within only 8 months. Cipeluch et al. (2010) chose five cities and towns to analyse the quality of OpenStreetMap in Ireland against Google Maps and Bing Maps. It was very interesting to see that there was no clear winner among the three mapping platforms from three aspects: spatial coverage, currency and ground-truth positional accuracy.

Previous studies gave a useful insight into the ways of assessing the quality of OSM data, especially the accuracy and completeness of linear features such as roads and walkways. Most importantly, they have showed that most regions of Europe are significantly developed, with good quality of OSM data. In some places, such as China, the application and quality of OSM data is still not yet clear, which causes great hesitation in some researchers about whether they can use this abundant data. Analysing and evaluating the current situation of OSM data in China is therefore fundamental work for many fields of research, and is the main focus of this chapter.

There were several evaluating elements on geographic information, but Goodchild (2008) claimed that perhaps completeness, a measure of the commission or omission of data, was one of the most significant aspects of geospatial data quality for VGI. He also raised the issue of positional accuracy for VGI as the product of untrained and unqualified volunteers. This chapter therefore focuses on two essential aspects of quality assessment – completeness and positional accuracy so as to outline the status of OSM data by comparing it with the Baidu data in China, followed by an analysis of its characteristics of development and distribution from 2011 to 2013.

Against this background, the first section picks up the quality of VGI data and provides a literature review for assessing the quality of OSM data. The following section describes the definition and status of two elements, completeness and positional accuracy, which were the main focus for assessing the quality of OSM data. The next section considers pre-processing in data sources which are acquired from OSM and referenced datasets in Chinese areas, and then describes in detail the methods for assessing the two elements and the development status of OSM data. Section 4 assesses results of completeness and positional accuracy, as well as the status of OSM data in China based on roads and point of interest (POI) features, via methods which are described in the last section. Finally, Section 5 draws conclusions.

2 Completeness and positional accuracy

The quality of geographic data was internationally standardised by the ISO/TC 211[2] Geographic Information / Geomatics Organisation in 2011. The International Standard (ISO 19113:2002) established five elements of data quality, completeness, logical consistency, positional accuracy, temporal accuracy and thematic accuracy, to describe the quality of geographic data and specify components for reporting quality information, and most importantly, to facilitate the comparison and selection of the dataset best suited to application needs or requirements. The organisation believed that it could encourage the sharing and use of appropriate datasets through complete descriptions of the quality of a dataset. For this reason, two elements of quality assessment are described here.

Goodchild (2008) put out a series of questions that communities needed to ask, and the first one is "What should positional accuracy assessment mean in a world in which everyone is a potential use of geospatial data?" The concern is perhaps greater now than in the past with the growing number of communities who lack a professional background and experience of geospatial data. This issue impacts confidence in using VGI as an alternative to authoritative datasets which are produced by more traditional means by commercial or government mapping agencies. It is therefore quite important to study the positional accuracy of OSM data collected by individuals with no experience or training, not associated with a brand and no set of official standards to follow. Positional accuracy is often measured through comparisons between a test dataset and one or more reference datasets which are assumed to represent the basic truth (Goodchild 1993). The methods outlined in Section 3 of this chapter will illustrate how to compute the accuracy of such data on points and lines.

Completeness, defined as a measure to describe the relationship between objects and the 'abstract universe' of all such objects, is as important as accuracy when it comes to assessing the quality of VGI (Van Oort 2006). It is therefore crucial to obtain knowledge of the existing number of features in a study area. Feature completeness can be defined over space, time, or theme (Veregin 1999). For example, a database is incomplete if it only depicts the locations of buildings in some areas, or only depicts residential buildings. Thus, completeness is always concerned with errors of "omission" which represent the absence of data. However, errors of "commission" for datasets may also occur, and represent the presence of excessive data (Devillers and Jeansoulin 2006).

---

[2] ISO/TC 211 Geographic Information/Geomatics. Http://www.isotc211.org/. Accessed 20 Mar 2013.

3 Materials and methods

3.1 Datasets used and pre-processing

As mentioned above, this chapter extended the method from previous studies which assess the quality of OSM data through comparing the OSM dataset against a reference dataset. As OSM is the test data, it was necessary to explain how OSM data was obtained and how it was pre-processed for analysis. Services, such as GeoFabrik and CloudMade, provided OSM data which could be downloaded by users for free in several formats such as OSM-XML and Shapefile (SHP), through an OSM database. An important characteristic of these downloaded files was that the OSM spatial data was close to real-time representations. GeoFabrik reported that "essentially any change made to the global OSM database is usually reflected in the data packages available for download within 24h" (download.geofabrik.de)[3]. These XML files contained none of the historical data, however. CloudMade leveraged rich OpenStreetMap data and completed it with various datasets from alternative sources for developers and application publishers with the most comprehensive and accessible location data, to support the most compelling Geo-enabled products. As big believers in community projects and crowd-sourcing, it has actively supported OpenStreetMap by providing public relations and legal support, sponsoring developer events and contributing to the OSM code base (cloudmade.com)[4].

All data used in the study was downloaded from GeoFabrik and CloudMade websites. The data within the Chinese boundary including three years of global data updated to 13 December 2011, 17 January 2013, and 22 February 2014, as well as local data, contained 31 provinces, 2 special administrative regions and data from Taiwan updated to 13 December 2011. Data was constructed as two point features: 'POI' and 'location'; three polyline features: 'roads', 'administrative' and 'coastline'; or two polygon features: 'natural' and 'water'. For the purpose of the comparison, all types of roads, such as primary, secondary, and cycleway, as well as POIs, that were hospitals, stations, and hotels etc. were used in this study, as roads are the core feature collected by OSM volunteers and POIs are most closely related to the daily life of humankind.

Haklay (2010a) noted that the selection of the reference dataset that was assumed to represent a version of reality of higher quality was a basic problem inherent in the quality assessment of any spatial dataset. Therefore, an authoritative dataset could be concerned to be the reference dataset. Baidu map is one of the most popular online maps in China, with abundant information, and is usually used for queries and navigation by ordinary people. As with OSM data, we will only compare roads and POIs. Further reference data was collected from the mileage of roads in the China Statistical Yearbooks of 2011 and 2012 (2013 has not yet been published).

---

[3] OpenStreetMap Data Extracts. Http://download.geofabrik.de/. Accessed 22 Feb 2014.

[4] CloudMade. Http://cloudmade.com/about. Accessed 5 Oct 2012.

This is one of the most authoritative books, with comprehensive and statistically exact information about the status of Chinese in most of fields, including transportation and population.

ArcGIS vision 10.0 software was used to pre-process the two datasets. The geographic data used in the OSM dataset was the WGS-84 lat/long data while in the Baidu dataset it was D_MAPINFO. In order to maintain consistency, we converted this data to the Universal Transverse Mercator (UTM) coordinate system zone 50 by using narrow zones 6° of longitude in width.

3.2 Global completeness

Although Steve Coast (2007), who established the OSM, noted that “it is important to let go of the concept of completeness”, many researchers have made progress in defining which areas are well covered and which are not (see Section 1). Here, two reference datasets, the China Statistical Yearbook and the Baidu dataset, have been used to test the completeness of OSM data. We computed the total length of ‘roads’, and then preliminarily estimated the global completeness of ‘roads’ compared to the actual length of roads reported in the China Statistical Yearbook. Table 14.1 show the results for both 2011 and 2012.

Table 14.1 Comparison of total length of ‘roads’ between OSM and China Statistical Yearbook

Year	Actual mileage (km)	OSM mileage (km)	Completeness (%)
2011	4106387	496296.2	12.1
2012	4193000	574357.9	13.7

From Table 14.1, it is very clear to see that the ‘roads’ data, the essence of OSM, was deficient in China, in 2011 and 2012. However, this result was too general to represent the completeness of OSM data in specific areas of China. It was worth paying attention to which areas involved ‘omission’ and which involved ‘commission’. Following Haklay’s (2010a) studies, the difference in data captured by OSM and Baidu map, the other reference source, was a core principle for assessing completeness. Here, all types of ‘roads’ were used in comparison instead of studying different levels of roads as Haklay did. This is because the classification standards of Baidu data differed to those of OSM data. For example, types such as primary and secondary roads in OSM data were classified based on traffic volume, while types in the Baidu data such as national, provincial and county roads were classified according to their administrative status.

In order to prepare the dataset for comparison, a 100 x 100 km^2 grid was created for the whole of China, to be divided into individual areas of completeness and deal with the VGI heterogeneity problem. The grids were then used to clip roads in the two datasets into segments along the grid lines. The lengths of segmented roads were then joined to each corresponding grid. Finally, the differences between two datasets were assessed using the following formula (Haklay 2010a):

$$\Sigma \text{(OSM roads lengths)} - \Sigma \text{(Baidu roads lengths)} \qquad (1)$$

3.3 Local completeness

After analysing the global completeness of OSM data, there should be a general cognition of the global status of OSM in China. However, local knowledge was not clear, which would be a problem when using OSM data in specific applications. As Girres and Touya (2010) found that one of the limitations for OSM data was heterogeneity, it was necessary to classify the distribution of OSM data between cities and other locations so as to quantitatively visualise the spatial heterogeneity. Local completeness was more variable than global completeness. Here, we determined a rule to demarcate the different levels of the quantity and density of OSM data which also could be regarded as a kind of completeness. All the ‘roads’ in polylines and points data from OSM were used in this.

3.3.1 Defining completeness levels

How to define and represent the different levels for completeness or incompleteness of data in locally? Haklay et al. (2010b) suggested that this notion of better quality needs to be evaluated quantitatively for different areas, which also required a clear definition. Accordingly, this chapter defined four levels of typical completeness (Table 14.2). Here, a part of Taka-Töölö, the oldest suburb of Helsinki - now considered by most a part of the central area[5], was selected as an example to represent the four different levels.

Table 14.2 Description of completeness levels.

Completeness levels	Examples
Most completeness -Areas contained most detailed map information. For instance, the points covered restaurants, eating and drinking places, airports, traffic lights, bus stations, telephone booths, toilets, ATMs, tourist spots, etc. The polylines had an abundance of roads including different levels of information and most of the primary roads were annotated with directions. Polygons gave an almost fully description for park, water, and building information. These areas could be applied to most purposes with satisfying.	

[5] Completeness example. Http://wiki.openstreetmap.org/wiki/Completeness_example. Accessed 6 Dec 2013.

Completeness - Most information but lacked a little ‘low-level’ data such as a part of POIs or footways.

Incompleteness - Areas had the basic polyline and polygon information but lacked a large amount of residential or POI data.

Most incompleteness - There was almost no data in these areas. Several main roads were included but without street names.

3.3.2 Demarcating completeness levels methods

More importantly, there were no authoritative answers about how to determine the classification intervals. This chapter provides a rule to quantitatively demarcate the different degrees of completeness of OSM data. The 100x100 km^2 grid which was used in assessing global completeness, could also be used at here.

Chinese territory covers approximately 9.6 million km^2. As the area of each grid was 10 000km^2, theoretically the total number of grids (entire grids) required to cover the whole territory of China should be 960. However, the actual number of grids, when counted, was exactly 1118, which was a little bigger than the theoretical grids because extra grids were involved around parts of the national boundary. These extra grids had no influence on the completeness of their positions, but they might affect the statistics for incompleteness. In order to reduce this impact, these grids should be removed when statistically analysing the incompleteness. We then quantitatively formulated the following classification rules for lines and points based on the number of theoretical grids.

- **Polylines density** (Table 14.3). Both averages of actual and OSM total length of roads per grid were divided by 960 that was the number of global grids. We defined those grids, whose values were larger than the average of actual mileage as belonging to complete groups, otherwise they would be assigned to incompleteness. Grids with values smaller than the average length of OSM roads would be classified to most incompleteness, which was to say, there was almost no data in these grid areas. As there were no statistics about the actual total length of roads in 2013, we therefore assumed the intervals in 2013 extending the same as in 2012.

Table 14.3 Rules of polylines density.

Year	Scale	Actual length (km)	OSM length (km)	Most Completeness	Completeness	Most incompleteness
2011	Global grids	4106387	496296.2	>6357	>4277	<520
	1 grid	4277.4	520			
2012	Global grids	4193000	574357.9	>6768	>4368	<600
	1 grid	4368	600			
2013	Global grids	-	1259120.0	>6768	>4368	<1311
	1 grid	-	1311			

- **Points density** (Table 14.4). According to polyline density, the actual data was nearly 10 times as much as the OSM data. We therefore assumed this ratio would emerge as similar in the points data. In other words, grids whose values were ten times bigger than the average quantities of OSM points would belong to complete groups. Grids with values smaller than the average amount of OSM points would be classified as the most incompleteness level.

Table 14.4 Rules of points density

Year	Scale	Number of OSM points (item)	Most completeness	Completeness	Most incompleteness
2011	Global grids	56625	>2400	>600	<60
	1 grid	60			
2012	Global grids	92304	>2400	>600	<60
	1 grid	60			
2013	Global grids	160377	>2400	>600	<60
	1 grid	60			

3.4 Positional accuracy

As with assessing completeness, the evaluation of the positional accuracy was carried out against the Baidu dataset since it contained a relatively high level of accurate information in China. In addition to the types of roads, the 'hospital' type in POI data was added as objects for comparison as they were public infrastructures which were a fundamental part of people's daily life. The methodology followed the study of Goodchild and Hunter (1997) by using buffers to determine the percentage of a line from the OSM dataset overlaying another higher accuracy dataset within a certain distance. Examples of lines and points are shown in Figure 14.2. As Haklay reported (2010a), OSM would ideally be treated as accurate within a region of about 20m from the true location, the *x* of the buffer for Baidu roads should be set at 20m, and, following Goodchild and Hunter's method (1997) the OSM dataset was buffered by 1m to calculate the overlap. The points used in this study were 'hospital' features. We assumed that if points were located in any locations within the areas of hospital coverage, such points could be treated as accurate. Here we supposed the *x* for the buffer of Baidu hospitals was 150m, that was to say, each point of the hospital features that was located within the circle buffer, with 150m, should be considered accurate data. It was important to note that the buffer for the points needed to be dissolved, as some adjacent hospital points from the Baidu data sometimes represented the same hospital, for example one referred to the eastern gate of a hospital while another represented the southern gate. If this procedure is ignored, the number of accurate points counted overestimated, which might influence the percentage of accurate data for both roads and hospitals in the subsequent analysis.

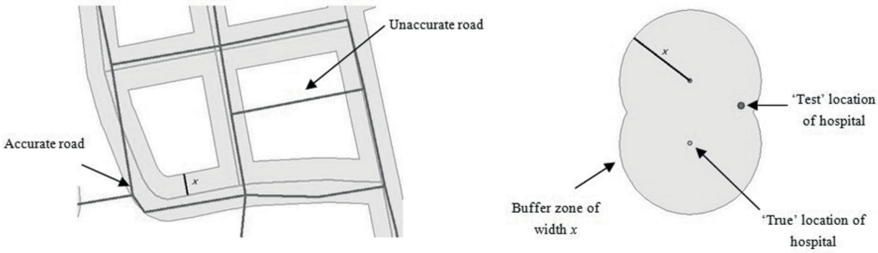


Fig. 14.2 Creation of buffers of width *x* for (a) lines (in this case roads); and (b) points (in this case hospitals); where red represented OSM data and green represented Baidu data.

4 Results and discussion

4.1 Applying the method

4.1.1 Completeness

All the roads in the two datasets, except those in the Taiwan area (since the Baidu dataset did not include this data) were used in assessing global completeness. The results in Figure 14.3 were calculated using Equation 1 as instructed in Section 3. It was clear that there was less data in OSM for most areas than there was in the Baidu data, which covered 71% of China, especially in the south eastern areas. More detailed OSM data existed mainly for the western cities and made up 24% of data for the whole country.

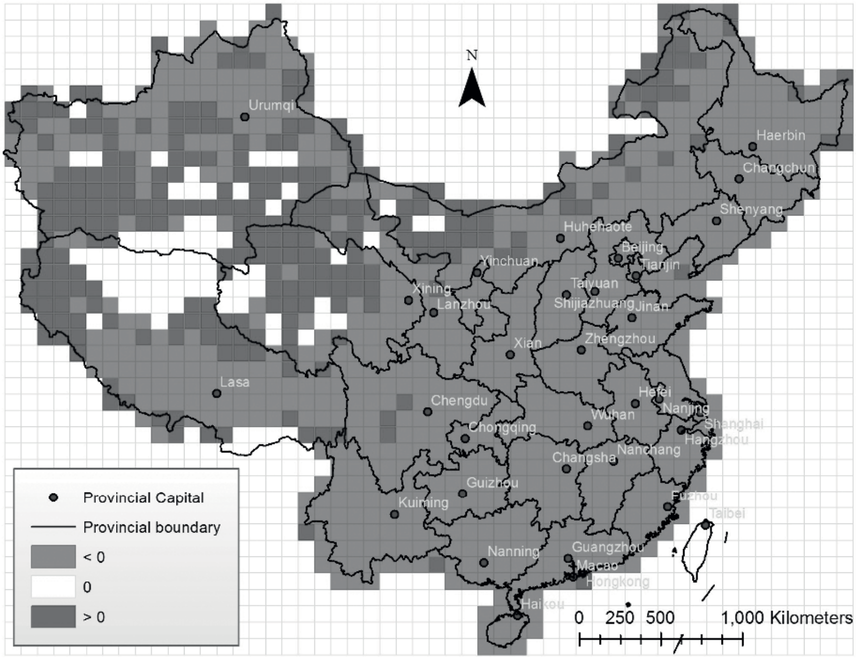
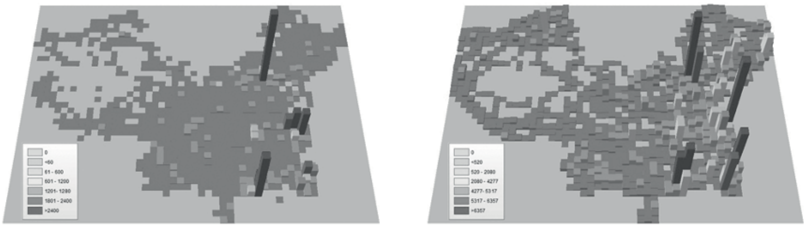
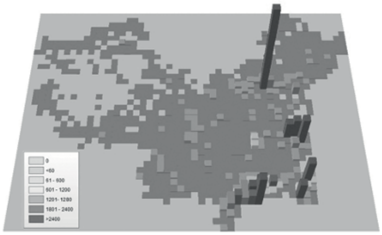


Fig. 14.3 Global completeness of OSM data in China, where red grids refer to OSM as more detailed than Baidu data, blue grids represented Baidu as more detailed than OSM data, and white grids indicated equal or no data.

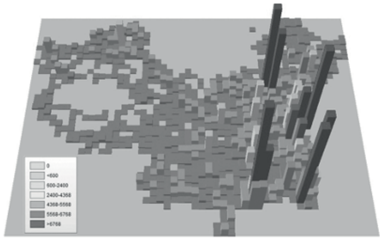
Based on the definition and classification rules which were introduced in the last section, the points and polylines of OSM data in China (including for Taiwan), from 2011 to 2013, were examined to classify the tiles into different levels of completeness. In order to increase the visibility of the gaps between completeness levels among cities, six results were finally extruded from 150 in ArcScene (Figure 14.4).



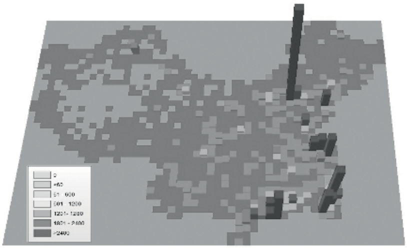
(a-1) Density of points in 2011 **(b-1)** Density of polyline in 2011



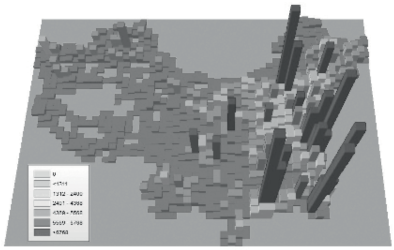
(a-2) Density of points in 2012



(b-2) Density of polyline in 2012



(a-3) Density of points in 2013



(b-3) Density of polyline in 2013

Fig. 14.4 The density of points and polylines using 100km grid squares. Grey means that 0 items fell into the corresponding grid which reflected a lack of data in the ten thousand square kilometre area. The redder the colour becomes in the grids, the more data has been contributed for these areas and the level of completeness of OSM data in China is higher. Conversely, the cooler the colour becomes, the lower the level of completeness given.

Figure 14.4 clearly depicts the different levels of the OSM dataset from 2011 to 2013 in China. The majority of points and polyline densities were green and some areas were even gray, which reflects the paucity of OSM data for China so far. The most abundant data for the points in 2011 (Figure 14.4(a-1)) is clearly from four cities: Beijing, Shanghai, Nanjing and Hong Kong with Beijing having the most. The four cities remained the principal abundant cities through 2012 and 2013 (Figure 14.4(a-2, a-3)). Some cities increased quickly, however, especially in Guangzhou, Taipei and Kaohsiung. The data from polylines was globally richer than that from points. Beijing and Shanghai also contributed the most polylines in 2011 (Figure 14.4(b-1)). Two regions, Guangzhou and Taiwan, developed so quickly that the richness of data in these areas was at the same level with Beijing and Shanghai in 2013 (Figure 14.4(b-2, b-3)). It was clear that the growth of OSM data in eastern cities was more and faster than in western cities of China.

Comparing the results of global completeness and local completeness (Figure 14.3 and 14.4), there appears to be a very interesting conflict between them at the first glance. Western areas had more detailed OSM data compared with Baidu data in Figure 14.3, while eastern cities possessed the richer data from OSM in Figure 14.4. Did these results really conflict, or is something wrong with the comparison? No. Normally, it is presumed that the areas with more abundant data are also more detailed and that it is impossible for poorer areas to contain more detailed data. However, sometimes this situation is reversed, as it is in this study. Although some areas in eastern cities are quite rich in OSM data, they were still poorer compared with the Baidu data which contained much more information than did OSM in these areas. In this case, some grids in Figure 14.3 luckily becomes red might caused by presence of only one line of OSM roads corresponding to where no line in Baidu dataset. OSM data in these areas therefore appeared as positive values. This phenomenon did not mean that western areas in China actually contained richer data than eastern cities, however, it reflected the advantage of OSM data in areas that are poor and economically underdeveloped.

4.1.2 Positional accuracy

Five study areas, Beijing, Shanghai, Urumqi, Guangzhou and Wuhan, were chosen to evaluate the positional accuracy of OSM in China, based on the methods introduced in the last section. According to the analysis of completeness, Beijing and Shanghai were the richest cities from 2011 to 2013, which was why these two cities were chosen to be sample cities for evaluating the quality of OSM data. Urumqi was a developing city located in the west of China which was used to represent a relatively poor area. Guangzhou is one of the fastest growing areas of OSM data according to Figure 14.4, and Wuhan, at where located one of the best university in the field of Geographic information system. The results of positional accuracy in five cities were displayed in Figure 14.5.

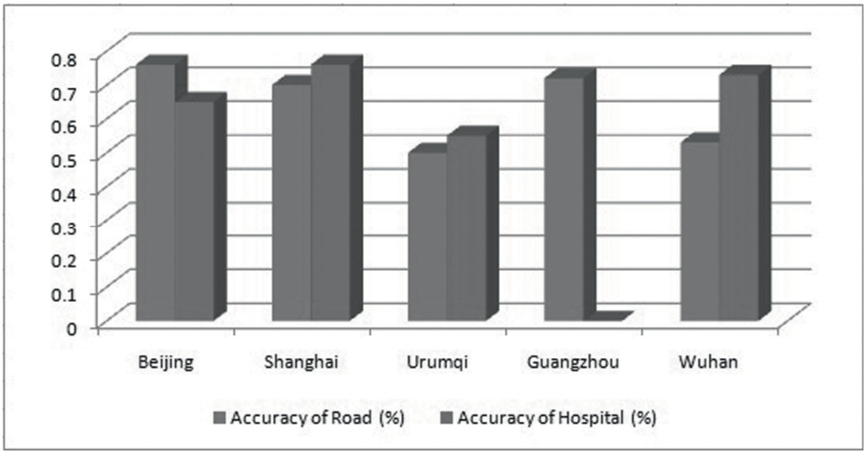


Fig. 14.5 The accuracy of roads and hospitals in five cases.

On the basis of the analysis, the OSM dataset provided a relatively good representation of roads and hospitals, with an average accuracy for roads and hospitals of nearly 66%, where the variation ranged from 50% to 76% (the accuracy of hospitals in Guangzhou was not counted in the statistics as there was no data in this area).

Beijing and Shanghai, two areas with the most completeness of OSM data, consistently had a higher accuracy for roads and hospitals compared with the other cities. Although the OSM data in Guangzhou grew quickly, the accuracy for hospitals was 0 which reduced confidence in the quality of OSM data in such areas. The quantity of OSM data in Wuhan was not outstanding, but its quality was quite high. In terms of Urumqi, not only was the accuracy of OSM data relatively lower than the other four cities, the amount of contributions was also low.

4.2 Analysing the growth of OSM data in China over 3 years

According to Figure 14.3 and 14.4, cool colours covered most of the Chinese territory. It was clear that the OSM data in China was not bountiful. Based on the actual number of grids, which including complete and extra grids, we counted the number of grids of different completeness levels (remove the extra data from most incompleteness levels) and calculated each growth rate over three years (shown in Table 14.5 and 14.6). Grids with the most completeness in points and polylines of OSM data covered less than 2% of China, and were formed of cities such as Beijing and Shanghai. According to the incompleteness regions in two statistical tables, more than 98% (98.23% for points and 99.17% for polylines) of the whole of China had incomplete OSM data in 2011. By the end of 2013, the ratio of incompleteness had decreased to 94.37% for points and 95.73% for polylines. This suggested that there has been a slightly positive improvement at each level except that of the most incompleteness for two years. The number of completeness grids has been growing whereas the number with the most incompleteness has been declining.

Table 14.5 Different levels of completeness for points data

	Area (10000 sq.km)			Percentage (%)			
level	2011	2012	2013	2011	2012	2013	Growth rate (%)
Most completeness	5	9	12	0.52	0.94	1.25	0.73
Completeness	12	29	42	1.25	3.02	4.38	3.13
Incompleteness	82	98	189	8.54	10.21	19.68	11.14
Most incompleteness	861	824	717	89.69	85.83	74.69	-15.00

Table 14.6 Different levels of completeness for polyline data

	Area (10000 sq.km)			Percentage (%)			
level	2011	2012	2013	2011	2012	2013	Growth rate (%)
Most completeness	5	8	19	0.52	0.83	1.98	1.46
Completeness	3	11	22	0.31	1.15	2.29	1.98
Incompleteness	170	205	247	17.71	21.35	25.73	8.02
Most incompleteness	782	736	672	81.46	76.67	70.00	-11.46

4.3 Characteristics of OSM data distribution in China

OSM mapping activity is closely related to many factors such as volunteer behaviours and environmental impacts. Thus, some social situations in China can be reflected through the distribution characteristics of OSM data.

- OSM data in eastern or capital cities is generally richer than in western or non-capital cities of China, which can be easily seen in Figure 14.4. Examples of these phenomena can be seen in Figures 14.6 and 14.7. These differences are probably caused by the economic gaps between eastern and western cities. As we know, western cities, which include more than half the poor population of China, fall a long way behind eastern cities in financial ability. Thus, the western cities have relatively lower need of accurate map data than the developed eastern cities.
- Contributions of OSM data for tourist resorts such as the Forbidden City in Beijing, or areas with a young population, are more abundant (Figure 14.8). Higher interests in a place or research absolutely attract more volunteers to contribute with their local knowledge. This is why places with much population movement, such as tourist attractions, bus or train stations and parks acquire more information from OSM users.
- Richer and more highly accurate OSM data appears in places where young people or researchers gather. Wuhan University, located in Hubei province, is apparently filled with abundant information (Figure 14.5 and Figure 14.9).
- Geographical environment is an important factor that cannot be ignored. It is clear to see in Figure 14.4 that there is a large area of continuous gray grid emerging from western areas. If you want to use any above impact factors to illustrate the reason why there is no data in these areas, it might not be enough. These gray areas correspond to the Taklimakan Desert, which occupies 330 000 km^2 and is the biggest desert in China. Virtually no people live in there, except for some archaeologists, which is why the contributions are so poor, even non-existent. However, according to the global completeness compared in Figure 14.3, there are great advantages in unprofessional volunteers creating and using detailled maps for specific purposes, which could give encouragement to those in the developing areas.

(6-a) Eastern **(6-b)** Western

Fig. 14.6 Samples of OSM in eastern (a) and western (b) cities: (a) in the Lujiazui area of Shanghai city where data is quite diverse and complete; and (b) in the Renmin Park area of Urumqi in Xinjiang which lacks the most information. Both are at the scale of 100 meters.

(7-a) Capital **(7-b)** Non-capital

Fig. 14.7 The difference of completeness in the capital city and a non-capital city: (a) in the area Urumqi of the capital city, Xinjiang, and (b) Changji, non-capital city of Xinjiang. Both are at the scale of 1000 meters.

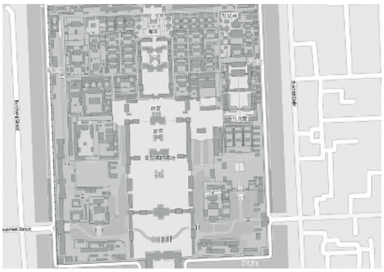


Fig. 14.8 Sample of resorts of the Forbidden City in Beijing

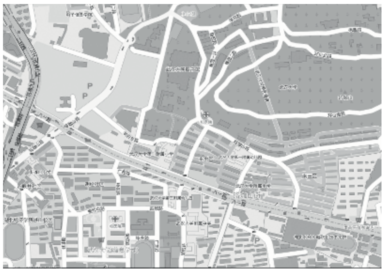


Fig. 14.9 Wuhan University campus in Hubei province

5 Conclusions

Since there has not yet been a clear and comprehensive study of the application status and quality of OSM data in China so far, researchers often hesitate about whether this popular data is suitable for their research purposes. It is therefore necessary to make an assessment of this. This chapter chose completeness and positional accuracy to evaluate the quality of OSM data in China from 2011 to 2013. Results showed that the OSM data in Beijing and Shanghai is mostly complete, with high positional accuracy. In addition, the quantity of OSM data has grown quickly in two years, especially in Guangzhou and Taiwan. However, the data is still extremely poor in terms of the whole country, as regions of incompleteness totalled more than 94%. Richer OSM data often appears in places with specific characteristics, such as developed eastern cities, capital cities, areas of population movement, and places of higher education. Although these are crucial factors, geographical and natural environments should not be ignored. The most interesting finding is that OSM has shown positive advantages for those who works in relatively poor cities.

This study gives an overview for researchers or governments who are considering the use of OSM data. Completeness and positional accuracy are essential parts of assessing the quality of OSM data, but other elements that determine the quality, such as logical consistency, temporal accuracy and thematic accuracy, also needed to be assessed. As most contributors are non-professionals, problems such as logical inconsistency, for instance features appear topological errors in a network may arise. Mooney and Corcoran (2012) found that some characteristics of "heavily edited" objects in OSM had been edited 15 or more times. These, of course, warrant further study regarding the quality of OSM data in China in the future.

Acknowledgments We thank OSM for the data used in this work. Some pictures, such as Figure 14.1 and Table 14.2, are OpenStreetMap copyright. We would like to express our appreciation to previous researchers, whose work inspired us. Moreover, we are very grateful to Yang Zhang and other colleagues in our laboratory for their valuable comments.

References

Cipeluch B, Jacob R, Mooney P (2010) Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps. In: Tate NJ and Fisher PF (eds.), Proceedings, Ninth International Symposium on Positional accuracy Assessment in Natural Resources and Environmental Sciences, Accuracy 2010, Leicester, UK, July 20-23, p 337-340.

Coast S (2007) OpenStreetMap. Workshop on Volunteered Geographic Information. Organized by NCGIA, Los Alamos National Laboratory, the Army Research Office and The Vespucci Initiative. Held in Santa Barbara on December 13-14.
URL: http://www.ncgia.ucsb.edu/projects/vgi/. Accessed 10 Oct 2013.

Codescu M, Horsinka G, Kutz O et al (2011) DO-ROAM: activity-oriented search and navigation with OpenStreetMap. In GeoSpatial Semantics. Springer Berlin Heidelberg, p 88-107.

Devillers R, Jeansoulin R (eds.) (2006) Fundamentals of spatial data quality. London: ISTE. URL: http://library.wur.nl/WebQuery/clc/1898206.

Girres JF, Touya G (2010) Quality assessment of the French OpenStreetMap dataset. Transactions in GIS, 14(4): 435-459.

Goodchild MF (1993) Data models and data quality: problems and prospects. Environmental Modeling with GIS, p 94-103.

Goodchild MF, Hunter GJ (1997) A simple positional accuracy measure for linear features. International Journal of Geographical Information Science, 11(3): 299-306.

Goodchild MF (2007) Citizens as voluntary sensors: Spatial data infrastructure in the World of Web 2.0. International Journal of Spatial Data Infrastructures Research, 2: 24-32.

Goodchild MF (2008) Positional accuracy 2.0. In Proceedings of the Eighth International Symposiumon Spatial Data Accuracy Assessment in Natural Resources and Environmental Sciences(ACCURACY'08), Shanghai, China, p 1-7.

Goodchild MF, Li L (2012) Assuring the quality of volunteered geographic information. Spatial Statistics, 1: 110-120.

Haklay M (2010a) How good is volunteered geographical information? a comparative study of OpenStreetMap and ordnance survey datasets. Environment and Planning B: Planning and Design, 37 (4): 682-703.

Haklay M, Basiouka S, Antoniou V et al (2010b) How many volunteers does it take to map an area well? The validity of Linus' Law to Volunteered Geographic Information. The Cartographic Journal, 47(4): 315-322(8).

Hentschel M, Wagner B (2010) Autonomous robot navigation based on OpenStreetMap geodata. In: 13th International IEEE Conference on Intelligent Transportation Systems, p 1645-1650.

Jacob R, Zheng JH, Ciepluch B, Mooney P et al (2009a) Campus guidance system for international conferences based on OpenStreetMap. In: J.D. Carswell, A.S. Fotheringham and G. McArdle (eds), Web and Wireless Geographical Information Systems, Proceedings. Lecture Notes in Computer Science. Springer-Verlag Berlin, Berlin, p 187-198.

Jacob R, Zheng JH, Winstanley AC et al (2009b) A multi-lingual pedestrian navigation and campus guidance system using CloudMade API. In: Proceedings of 10th International Conference on GeoComputation, Sydney, Australia, p 473-477.

Mooney P, Corcoran P (2012) Characteristics of heavily edited objects in OpenStreetMap. Future Internet, 4(4): 285-305.

Neis P, Singler P, Zipf A (2010) Collaborative mapping and Emergency Routing for Disaster Logistics - Case studies from the Haiti earthquake and the UN portal for Afrika. In Car A. et al. (eds.): Geospatial Crossroads @ GI_Forum 2010. Proceedings of the Geoinformatics Forum Salzburg, p 239-248.

Over M, Schilling A, Neubauer S et al (2010) Generating web-based 3D city models from OpenStreetMap: The current situation in Germany. Computers, Environment and Urban Systems, 34(6):496-507.

Ramm F, Topf J, Chilton S (2011) OpenStreetMap: Using and Enhancing the Free Map of the World. UIT Cambridge, UK.

Van Oort P (2006) Spatial Data Quality: From Description to Application. PhD thesis, Wageningen Universiteit. Published by the Nederlandse Commissie voor Geodesie (Netherlands Geodetic Commission).

Veregin H (1999) Data quality parameters. In P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind (eds), Geographical Information Systems, 2nd Ed., Chap. 12. New York, NY: Wiley, 1:177-189.

Yin JJ, Carswell JD (2011) Touch2Query enabled mobile devices: A case study using OpenStreetMap and iPhone. In: 10th International Symposium on Web and Wireless Geographical Information Systems (W2GIS2011), Kyoto, Japan, Springer LNCS 6574: 203-218.

Zheng JH, Chen XY, Ciepłuch B, Winstanley AC, Mooney P, and Jacob R (2010) Mobile routing services for small towns using CloudMade API and OpenStreetMap. The International Archives of the Photogrammetry, Remote Sensing and Spatial Sciences, 38, Part 2.

Zielstra D, Zipf A (2010) A comparative study of proprietary geodata and volunteered geographic information for Germany. In: 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal, p 1-15.

Combining Geographical and Social Dynamics in Dynamic 3D Environments

Mingyuan Hu[1,2], Hui Lin*[,1,2,3], Weitao Che[2], Tianpeng Lin[2], Zhang Fan[2]

Abstract Mapping our real world is a challenging task, mainly because it involves numerous dynamics within the contexts of physical and social environments. A dynamic map environment was developed in this chapter, to help us understand and manage our complex world in a more sustainable and resilient way. A conceptual architecture to integrate real environmental information and dynamic social activities was presented as three elements: 1) based on a collaborative virtual geographic environment with volunteered knowledge construction; 2) linking the dynamic environmental factors that are relevant to our everyday life; and 3) expanding the comprehensive social experience/activities of daily interactions and communications without the limitation of spatial and temporal constraints. Using the virtual campus map of the Chinese University of Hong Kong, we verify this hypothesis and present our initial design of an interactive dynamic environment. The results show that the dynamic map environment can give people visual and tangible access to real-time information about environmental dynamics, and help them to join in and interact with actual happenings, from their daily activities to purposeful cognition of the world. The research provides an insight into the development of a dynamic map environment, and plays an important role in developing an open research platform for virtual geography and campus-based social cognition.

M. Hu, H. Lin
Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, 518051, China
Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong

W. Che, T. Lin, Z. Fan
Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_15,

1 Introduction

Using map media to cognize a whole living environment is a challenging task, mainly because it involves not only the geographical, but also the social space, which has historically been linked to various domains of everyday activity. The development of tightly integrated natural and social dynamics is still in the beginning stages and more research is needed to make progress in this respect (Sui and Goodchild 2011). A dynamic map environment is anticipated as being advantageous in cognizing our everyday lives because its emphasis puts the traditional map in a larger context, both spatially and socially, by involving more dynamic socio-spatial knowledge. It has the ability to facilitate daily life and to explain what is happening and how these actual happenings affect the world around us.

Research and development efforts in many domains, such as cartographical mapping (Friedmannová and Stanek 2010, Konecny 2011), geographical information systems (ArcGIS 2013), online digital environments (Butler 2006, Sheppard and Cizek 2009) and virtual environments (Montello et al. 2004), provide the support for understanding our living environments. The widespread use of computer technologies and access to wider information is driving mapping media away from static paper maps towards more interactive, dynamic and animated geographic representations, making it easier for people to understand our complex human-centred system (Konecny 2011). The availability of the GeoWeb (Elwood 2011, Scharl 2007) and information and communication technologies (ICTs) (Yang et al. 2009) stimulates an increasing online trend towards loosening the traditionally strong links between activity, place, and time, and naturally accommodates activity in cyberspace as well as in physical space (Haklay et al. 2008, Couclelis 2009, Hall et al. 2010). The use of online digital environments, such as digital globes and online mapping systems, has been seen as spatially-enabled information sharing and volunteer knowledge presentation based on large, complex data sets. For example, NASA's World-wind and the Google Maps' platform have already witnessed many innovative ways of gathering diverse information and sharing it in a visually appealing way, giving people the ability to progress from analysing static data to engaging with a dynamic spatial process of time-critical or real-time monitoring and decision-making (Dodge and Perkins 2009). In addition to conveying the dynamic heterogeneity of the geographical space and various spatial relations, the online map media are urged to integrate more human activities and observational human behaviour with the contributions of multiple participants. Completing the chain that links the social rules to the geography in the map could facilitate understanding of socially interacting systems (Goodchild 2012, Yu and Peuquet 2009, Miller 2003). As social media has become more locationally aware and people's experiences with their environment are more harmonised, the convergence of GIS and location-based social media has made it possible to explore dynamic social behaviour (Shaw and Yu 2009), focusing on the complex interactions between people, spaces, and places. The Copenhagen wheel project (Outram et al. 2010) can be seen as part of the general trend of inserting intelligence into our everyday lives to tap into the potential brought by the digital networking of people, objects and cities through real-time information loops (Resch et al. 2013). Furthermore, people have discussed for many years the possibility of having a people-oriented representation such as an avatar on Google Earth in order to augment their context-sensitive interactions (http://www.technologyreview.com/Infotech/18888/), especially for supporting the individual dimensions of an experience with strong affective connotations, social relationships, and feelings about the overall experience of actual happenings.

At the same time, sociable 3D virtual worlds (e.g. Second Life) and computer gaming engines, (e.g. Unity and Unreal) provide us with the dramatic potential to change and facilitate how people interact, access entertainment, and conduct business (Bainbridge 2007, Kreijns et al. 2007, Messinger et al. 2009). Some virtual worlds have been associated with imaginary scenes and game-like learning, which often makes the participants unreal or unfamiliar to their normal perception, leading to a dubious or optional feeling about this kind of anonymous world (Tuzun 2009). Although there is evidence supporting the fact that individuals do not require exact replications of what they are observing (Johnson et al. 2010), certain geographical scenarios have been employed in social behaviour research on virtual worlds to produce more reliable results (Brandão and Coelho 2011, Gilbert 2011). Pinpointing reasonable requirements in the virtual world that is equipped with familiar geographical scenarios makes interpretation of the dynamic environment more active (Mania et al. 2010, Ragan et al. 2012). Similarly, Konecny (2011) highlighted the human-centred virtual geographic environments (Lin and Zhu 2005, Lin et al. 2013a, Lin et al. 2013b) that is built with real geographical scenarios. It can inspire the development of cartography in connection with geographical information science and revolutionise the representation of the natural world. People's cognitive experiences of their living environments thus have to be extended from 2D static representations to networked virtual/digital reality in 3D virtual worlds and virtual geographical environments for dynamic geographic and social cognition (Huang et al. 2001, Kraak 2004, Djorgovski et al. 2009, Marchant 2010).

In the light of the abovementioned background, the present moment marks a major historical transition. More attention is being paid to opening up the social space by linking more realistic environmental effects (Ravitz et al. 2010), social networks (Corten 2012, Farman 2010), and daily spatial social behaviours (Blascovich et al 2002, NRC 2007, Pelachaud et al. 2007). Such a space will give the public a better understanding of geographical and social experience (Wood et al. 2010), allowing the participants to expand their abilities to carry out daily activities both in virtual world and in physical space (Kohler et al. 2009, Lifton and Paradiso 2012; Cook and Hopkins 2007). The links between the use of social media and mapping strategies across multiple disciplines (Crampton 2009) are leading towards a focus on the use of maps as a dynamic interface to rich geographical understanding and human-centred natural interactions (Meng 2010, Wilson et al. 2010).

The remainder of this chapter is organised as follows. The methodology for the dynamic map environment is proposed in Section 2. Section 3 illustrates the implementation issues and case demonstration of the dynamic campus map of the Chinese University of Hong Kong. Finally, Section 4 offers the conclusion to this work, together with thoughts on future prospects.

2 Conceptual Architecture: Towards the Dynamic Map Environment

In response to the changing role of map media, this paper proposes a conceptual framework to integrate real environmental information and diverse social activities into a dynamic map environment with the aim of providing an intimate tie between these two components (Figure 15.1). Using a dynamic map environment can enhance and facilitate human cognitive awareness relevant to dynamic natural and social environments, and provide a comprehensive social experience at scales from personal activities to socio-spatial practices. The conceptual framework consists of three levels (Figure 15.1): (1) constructing dynamic geographical scene, (2) merging dynamic environmental information, and (3) involving dynamic social activities.

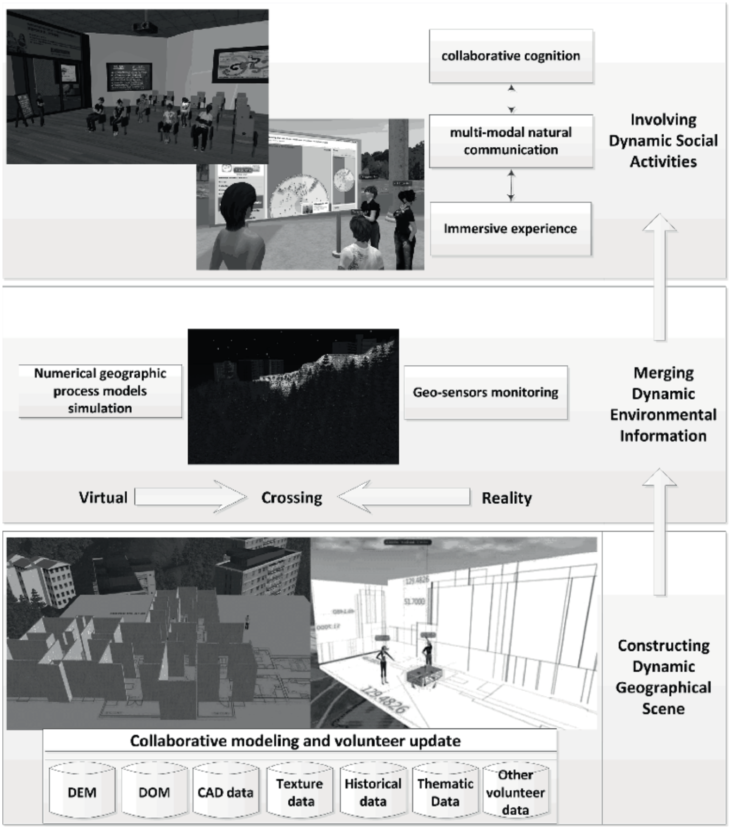


Fig. 15.1 The conceptual architecture of the dynamic map environment.

Constructing a Dynamic Geographical Scene: The dynamic map environment is essentially geographical, composed of a series of thematic data from our physical world. It allows the creation of both the physical and human geographic environments and provides users with a similar life experience with horizontal locations, 3D vertical structures and multi-dimensional geographical scene expressions. The dynamic scene is developed collaboratively and updated by volunteers. Firstly, collaborative modelling enables multiple remote modellers to construct the virtual environment using a basic foundation for the dynamic map. The full view of the scene is supported in hierarchy - from the simple bounding appearance to real 3D interior layouts. Meanwhile, we use the concept of Volunteered Geographic Information (VGI) (Goodchild 2010) by encouraging users to collect data voluntarily to update the virtual environment, so as to reduce the heavy costs of updating. A large voluntary collection of temporal data in different periods will provide an opportunity to build a past, present and planned scene, which can help people to understand the speed of changes, identify what happened in the past, and contribute to reliable future plans. It provides an experience of belonging with memorized things in a geographical space that would be beneficial for historical social researchers.

Merging Dynamic Environmental Information: The real world is in flux and consists of various dynamic phenomena, and many dynamic phenomena cannot be directly observed in reality (Hirtle and Macheachren 1998). A dynamic map should contain changing environmental factors to enhance or augment the human cognition of the dynamic process, as well as to help people become better aware of their living conditions. For example, comparing historic and real-time environmental data would help people to become aware of important trends in their living environment. Visually experiencing the diffusion processes of air pollution would help people to intuitively compare the pollution situations at different periods in a day. As the trend of "merging networked sensors and virtual reality" becomes more common in our daily life (Lifton and Paradiso 2012), innovative media technologies are changing the way in which people interact with each other and their environments. Geographical multi-source sensors thus form the most critical foundation of a dynamic map. The real-time monitoring data acquired from distributed sensors, such as wind speed and air pollution conditions, can be integrated into, and then visualised in the dynamic geographical environment. It further enhances our ability to understand the geo-environmental process and rules. In order to estimate or predict the conditions of the dynamic environmental factors, we integrate numerical geographical process models into the virtual geographic environment and provide users an easy-to-use interface to operate these functions. For example, the Computational Fluid Dynamics (CFD) model could make the transmission and dispersion of air pollution easier for the public to understand and further assist officers in decision making (Epstein et al. 2011).

Involving Dynamic Social Activities: On the basis of the underlying platform and dynamic environmental information display, we introduce the typical social features of our real world to the dynamic environment, which allow users to interact with each other virtually in social activities. At this level, avatar-based approaches are adopted to model online users' responses to dynamic activities and social communications, in consideration of their advantage in representing the combination of humans in the real world and 3D avatars in the virtual world. The dynamic map environment is thus moving towards becoming both a physical and virtual world, in which people can congregate to experience a true sense of community.

(1) The essential level of the social space is the experience by means of 3D virtual avatars, which depict the identities of online participants in the real world. While avatars are digital entities, their interactions and relationships are ultimately dictated by the thoughts and feelings of a physical person. The participants can 'step' into the virtual environment to acquire a great deal of spatial knowledge about environments, gain an immersive experience of being in the dynamic environment, thus acquiring an intuitive perception of the geographic environment as well as the natural phenomena around them.

(2) Secondly, it is important to realise that spatial properties of a directly experienced environment are not just "seen". They are sensed multimodally: the participants could engage in a multi-modal natural communication and collaboration that is very similar to that of human beings – information could be exchanged through visual, auditory, tactile, kinaesthetic, sixth sense and other means. It surpasses the limitations of the experience of textual and report type information content, providing a shared space to support multiple participants all over the world in meeting and collaborating simultaneously. The social space is able to move towards a much broader, more capable and timely information acquisition and management mode to provide participatory solutions beyond the isolated technology. This multi-modal communication can be customised to suit specific needs, for example, students and teachers can communicate/participate in interviews in a relaxed, memorable learning environment, which could promote much greater retention than reading a textbook.

(3) The third characteristic is combining individual and group behaviour. The dynamic map can support collaborative cognition, such as collaborative activity of urban planning involving urban geographers and voluntary members of the urban community who share some commonality with regard to their specialties, concerns and requirements. This kind of purposeful cognition tends to relate to geographical, social, and psychological elements. In the case of campus planning, the cultural value, ecological concerns and the sense of togetherness with students, staff and alumni can assist in determining potential development sites within the campus and what a campus will become.

3 Case Studies and Implementation Issues

In order to verify the methodology, a case study of the virtual campus map of the Chinese University of Hong Kong (CUHK) was designed and implemented as a dynamic campus environment. The implementation of the dynamic campus map of the CUHK was based on an open source, multi-platform, multi-user 3D application server OpenSim 0.7 (http://opensimulator.org/wiki/Main_Page), in which distributed data management, 3D visualisations, and interoperability can be flexibly realised, and a dynamic campus environment can be well developed. The development and implementation environments are:

Data Region Server: Dell (TM) PowerEdge (TM) R610 Rack Mount Server Intel(R) Quad Core X5550 Xeon (R) CPU, 2.66 Ghz, 8M Cache, 6.40 GT/s QPT, Turbo, 8GB (4X2GB) Memory.

Operating System: MS Windows Server 2008 and Ubuntu 10.04.

Programming Tools□Visual Studio 2010, Python 2.7, OpenSim0.7; Graphical Card: NVIDIA Quadro® FX 4800; Graphical Interfaces: OpenGL.

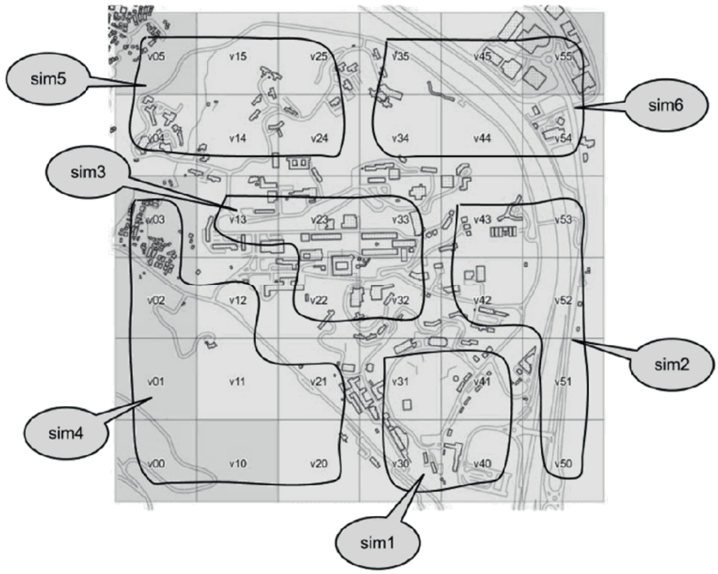


Fig. 15.2 The whole terrain domain of the CUHK campus and its distributed regions.

As shown in Figure 15.2, the whole terrain domain of the CUHK campus was divided into 36 regions by 6 distributed region servers and named from Sim1 to Sim6. Each region was further assigned a unique ID for a grid index such as v23, v31, etc. Based on OpenSim, the region server is a feature for serving one or more scene regions. It communicates with the grid manager and runs multiple region simulators on distributed machines with the ability to offer the easy potential to leverage the distributed data management and respective dynamic data balance. Based on synchronous or asynchronous collaborative 3D modelling, the entire virtual campus environment was constructed, from the simple bounding appearance to real 3D interior layouts, as shown in Figure 15.3. Specifically, the following modelling methods were oriented to multiple levels of detail (LOD) for 3D campus content creation: primitive-based modelling for quick 3D reconstruction with basic primitives; integrating the existing CAD to reduce the amount of effort and time required to create 3D virtual world contents and improve the level of detail of 3D objects. The sculpted method, which was employed for complex surface modelling, allows the creation of smooth, curved surfaces such as boats, sculptures or any other object from the absolute textures (for more detail on the basic construction process of the virtual CUHK campus, please see Hu et al. (2011).

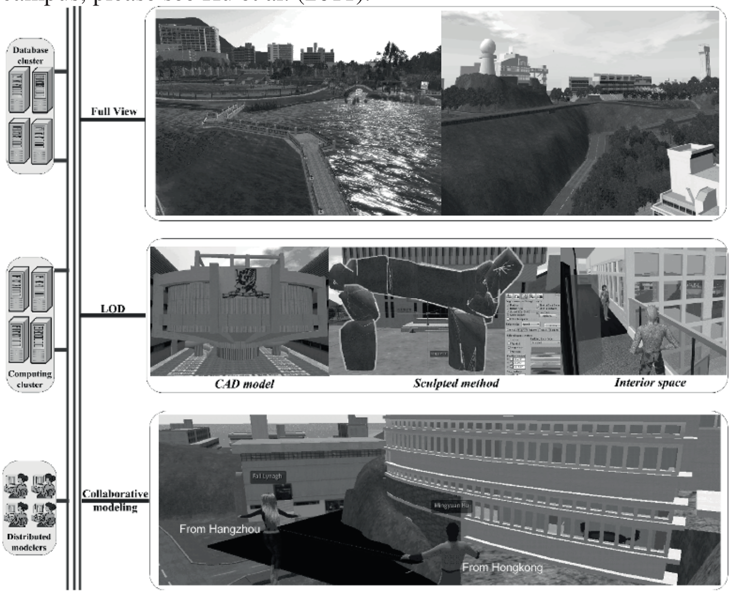


Fig. 15.3 Dynamic and volunteered construction of geographic scene with multi-user collaborative modelling.

On the other hand, according to the dynamic construction of the distributed campus geographic scene, we cannot entirely neglect the issues arising from the use of various volunteered information. Because the VGI information is not always exactly correct or reasonable, full control of the quality of this information requires the active involvement of scientists from both social and other relevant domains. In our case, the core team (distributed team) is responsible for the creation and maintenance of the main campus scene during the phase of the initial construction work, but due to the dynamic nature of updated data or dynamic information publishing, the input of volunteer users is greatly welcomed and appreciated.

Based on the highly realistic foundation of the geographical campus environment, the following examples show how a human-centred environment may merge dynamic environmental information and social behaviour into a virtual geographic environment to facilitate our daily campus life.

3.1 Merging dynamic environmental knowledge into a virtual geographic environment

To facilitate our perception of the everyday natural environment on the campus scale and further identify how actual happenings affect the campus environment around us, the dynamic campus map should involve much more real-time situation monitoring and assessment, and require tools that treat information as continually changing. In the context of dynamic campus environmental cognition, location-based SO2 concentration sensors have been distributed across the CUHK campus along the main residence roads to monitor the physical and environmental conditions. Together with these fixed sensor nodes as dynamic boundary conditions for small-scale numerical simulation, the dynamic concentrations of air pollutants were calculated for the local campus area. The whole computational process was solved using three-dimensional Navier-Stokes equations with a standard k-epsilon model for the incompressible flow of Newtonian fluids, and the species transport method was selected for air and sulphur dioxide mixing and transport to estimate and predict the conditions of traffic-related air pollution. As shown in Figure 15.4, the dynamic environmental factors of SO2 concentration were integrated and then visualised in a geographical scenario using various colours at discrete points. Moreover, virtual humans in the virtual space are obviously the properties of the virtual campus map, so avatar-based humans were designed and implemented in the campus map. These virtual avatars are able to explore and interact with each other within a three-dimensional environmental space. For example, they can dynamically experience the dynamic air pollution conditions along the road and buildings, and can therefore intuitively compare the different situations in the morning and evening.

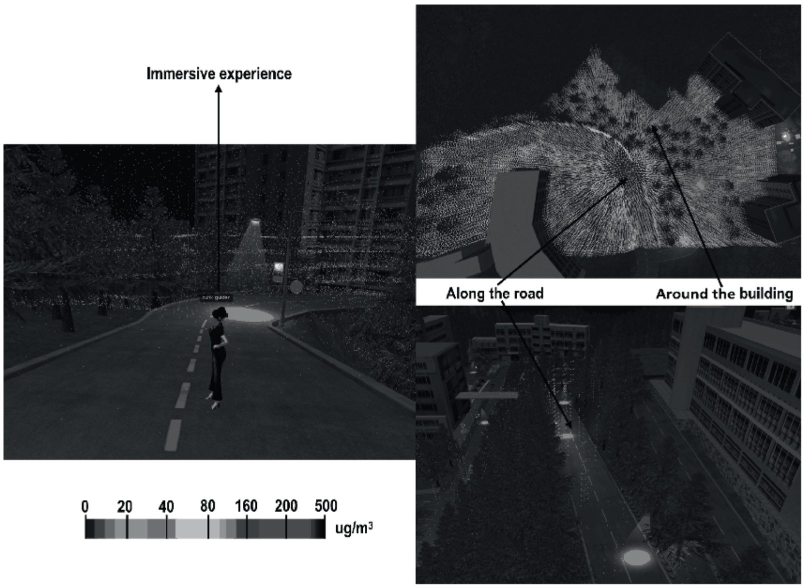


Fig. 15.4 Real time modelling and visualization of dynamic environmental information.

3.2 Merging certain social rules into virtual geographic environments to facilitate our daily life and activities

There are many daily activities that take place on campus involving certain social behaviour, so the dynamic campus environment is expected to play an irreplaceable role in many ways. The following hierarchical social space outlines how a dynamic map environment may represent and simulate perceived campus environments, and allow distributed, multi-participants to communicate and interact with each other; it also allows the implementation of purposeful social cognition.

(1) Firstly, in consideration of students, staff and alumni on the wider scale by collecting dynamic information about campus activities, the virtual campus environment can serve users as a dynamic information publishing and real-time feedback platform in an immersive 3D environment. As shown in Figure 15.5, dynamic information about daily shuttle bus services with their geo-referenced locations was shared in the virtual campus environment. Figure 15.6 indicates the initial repair information sent from our lab to an Estate Management Office (EMO) worker, showing the specific position of a malfunctioning facility and its building. Furthermore, by combining camera sensor data, such as real-time images or video transmission from a mobile, the EMO worker could quickly learn the extent of the malfunction, and what additional tools were required to repair the broken facilities before arriving at the real place, in order to deal with them. Important feedback can be actively obtained or answered before, during or after the repair process. It is obvious that the dynamic campus map has the potential to facilitate the daily affairs of the campus and improve the problem finding and problem solving skills of workers.

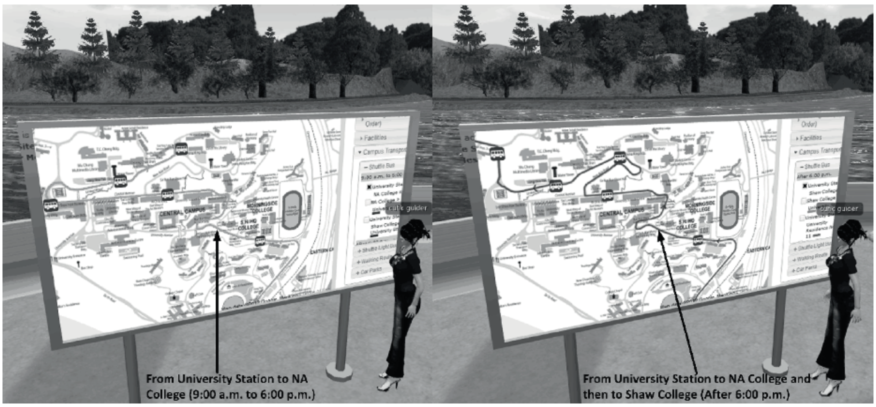


Fig. 15.5 Dynamic information publishing for daily shuttle bus services.

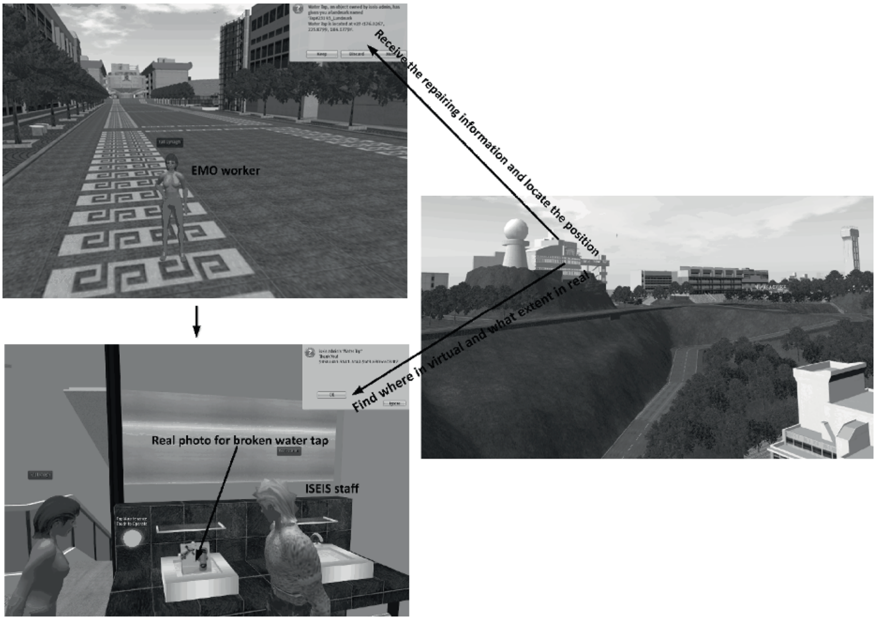


Fig. 15.6 Dynamic information feedback of EMO repair work: integrating real into virtual environment.

(2) In addition to the publication of dynamic information, Figure 15.7 shows the possibility of affecting the feeling of social presence and group awareness in collaborative learning situations. Based on various kinds of verbal and nonverbal communications, such as the virtual PowerPoint medium, real-time videos, and particular gestures, a student or a scientist could attend a virtual class, lecture, informal group discussion, or even full-sized international conference in a virtual educational environment. They could carry out exchanges of opinion simultaneously, without any need to travel in the physical world. This shift to a hybrid learning environment could potentially enable users to feel togetherness with others, especially those people with mobility impairment who are often isolated from society due to their lack of social contact with the outside world.

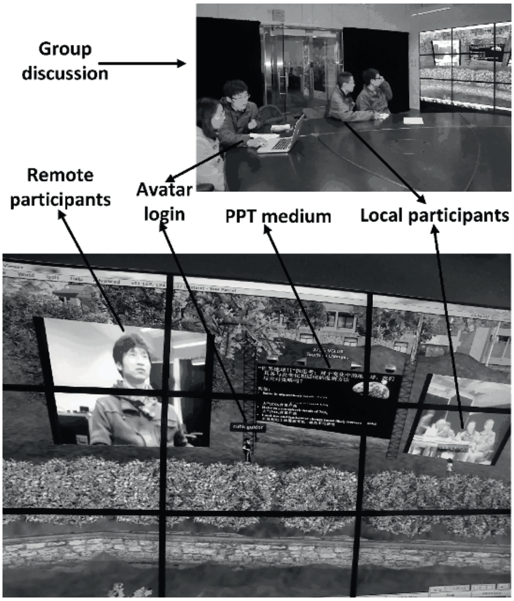


Fig. 15.7 Multi-modal interactions and collaborative learning situations in a shared space

(3) Typically, a familiar and dynamic campus map has been concerned with the requirements of the Campus Development Office (CDO) for timely decision-making with collaborative cognition by professional designers, teachers, students and alumni.

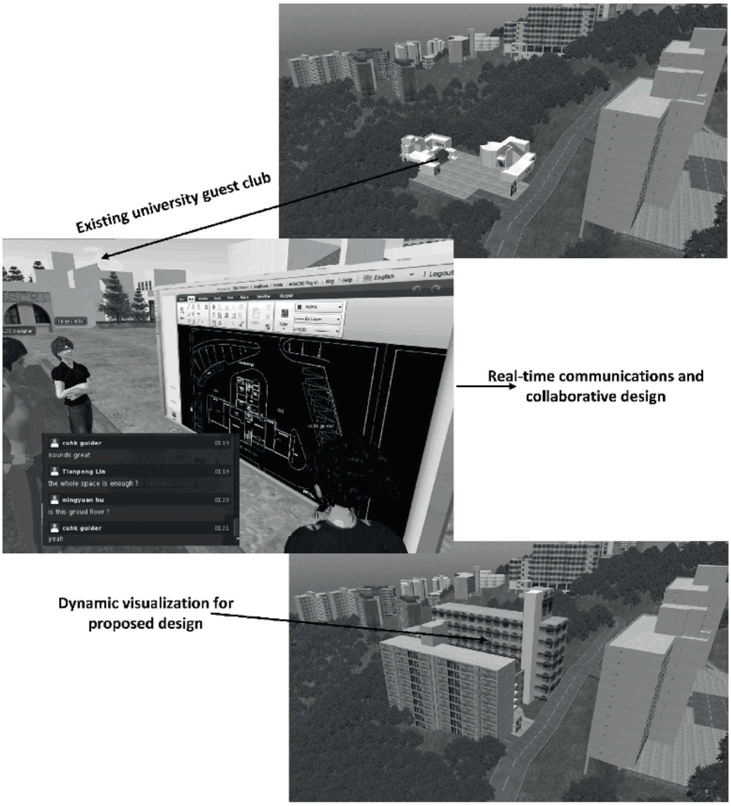


Fig. 15.8 Group communications, dynamic feedbacks and visualised experiences about proposed designs.

As shown in Figure 15.8, the online AutoCAD tools (https://www.autocadws.com) were integrated into the dynamic campus environment, which aim to facilitate professional designers and CDO workers in collaborative design with an embedded web browser. Meanwhile, proposed building designs were visualised alternatively in the same space with surrounding facilities and natural situations. Within this shared dynamic campus environment, other concerned groups, such as students, staff and alumni, were given the opportunity to provide their feedback and real-time communications about campus planning. It therefore becomes clear that the dynamic cognitive environment goes beyond visualisation to provide an optimised selection for the virtual experience and future design of better campus planning.

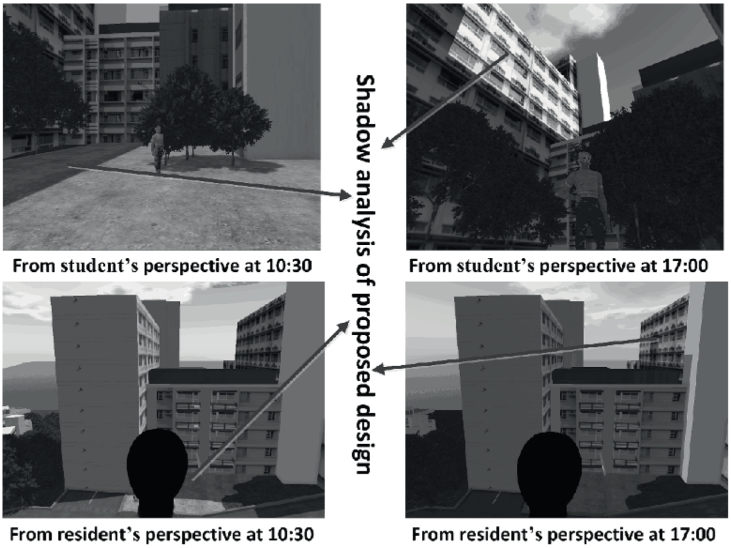


Fig. 15.9 Collaborative cognition from the resident's and student's perspectives on the proposed design of student hostels.

By combining the cognition experiments of different groups on a proposed design, it becomes clear that the dynamic campus map goes beyond visualisation to provide an optimised context and model for the virtual experience and future design of a better campus environment. For example, from the resident's perspective, a planned building can be viewed, using the system to assess the potential impact on their existing living conditions, and the apartment, such as the view from their apartment and the lighting conditions. A student can enter the newly planned building to experience its internal environment, lighting and other environmental conditions (Figure 15.9). When a student is in the central rest areas, the sunlight may be blocked by the east, or west wings and the student might have less time with sunlight shining into those areas between 10:30 am to 5:00 pm. At this level, in addition to the professional design and geographical environment, the avatars are active participants with the necessary psychological behaviour to achieve certain tasks in a specific experiment, thus letting humans generate a feeling of the "visual cognition process" of the real environment, improving their understanding and cognition of the real campus environment.

4 Conclusions and Future Research

Aimed at enhancing and facilitating human cognitive awareness relevant to dynamic lives at scales from personal activities to socio-spatial practices, a dynamic map environment was developed with three potential roles: 1) based on a collaborative virtual geographic environment with volunteered knowledge construction; 2) linking the dynamic environmental factors that are relevant to our everyday life; and 3) providing a comprehensive social experience of daily interactions and communications. Practically, a campus-based dynamic environment can be extended to a larger scale and embrace more social space by linking individual behaviour and social interactions. It expands our ability to represent and analyse the increasing variety of activities in space and time. Such research can inform us about how to construct a dynamic or live map environment, which preserves more of the social elements of real-life interactions and facilitates effective and efficient facility management, as well as the need for the digital protection of the university's historical and cultural relics.

This research provides dynamic information visualisation on environmental factors, but remains a challenge in dynamic assessment by integrating the real-time monitoring data acquired from distributed sensors with numerical geographical process models. This integration issue will be considered in future work. In addition, a valid solution to control the quality of VGI information needs to be provided to support the dynamic construction of the distributed campus geographic scene. Another important phase of this research will be focused on campus-based social research into cognitive psychological experiments and dynamic behaviour simulation, to build a dynamic cognitive environment that will facilitate better exploration of the human behaviour features that are difficult to fully understand via the physical campus.

Acknowledgments The work is supported by the National Natural Science Foundation of China (grant no. 41101370, 41371388 and 41171146) and the Open Research Fund of Key Laboratory of Disaster Reduction and Emergency Response Engineering of the Ministry of Civil Affairs under grant no. LDRERE20120302.

References

ArcGIS (2013) Mapping & Analysis for Understanding Our World. http://www.esri.com/software/arcgis/index.html. Retrieved 16 October 2013.

Bainbridge W (2007) The scientific research potential of virtual worlds. Science. 317, 472.

Blascovich J, Loomis J, Beall AC, Swinth KR, Hoyt CL (2002) Immersive virtual environment technology as a methodological tool for social psychology. Psychological Inquiry. 13, 103.

Brandão SP, Coelho A (2011) Procedural modeling for realistic virtual worlds development. Journal of Virtual Worlds Research. 4, 1.

Butler D (2006) The web-wide world. Nature. 439, 776.

Cook T, Hopkins L (2007) Social media: An introduction to the power of "web2.0". http://trevorcook.typepad.com/weblog/files/cookhopkins-socialmediawhitepaper-2007.pdf. Retrieved 01 September 2013.

Corten R (2012) Composition and structure of a large online social network in the Netherlands. PLoS ONE. 7, e34760. doi:10.1371/journal.pone.0034760.

Couclelis H (2009) Rethinking time geography in the information age. Environment and Planning A. 41, 1556.

Crampton J (2009) Cartography: Maps 2.0. Progress in Human Geography. 33, 91.

Djorgovski SG, Hut P, McMillan S, Vesperini E, Knop R (2009) Exploring the use of virtual worlds as a scientific research platform: The meta-institute for computational astrophysics (MICA). Proceedings of Facets of Virtual Environments, ICST Lecture Notes Ser. Berlin/Verlag: Springer. 29.

Dodge M, Perkins C (2009) The 'view from nowhere'? Spatial politics and cultural significance of high resolution satellite imagery. Geoforum. 40, 497.

Elwood S (2011) Geographic information science: Visualization, visual methods, and the geoweb. Progress in Human Geography. 35, 401.

Epstein JM, Pankajakshan R, Hammond RA (2011) Combining computational fluid dynamics and agent-based modeling: A new approach to evacuation planning. PLoS ONE. 6, e20139. doi:10.1371/journal.pone.0020139.

Farman J (2010) Mapping the digital empire: Google earth and the process of postmodern cartography. New Media & Society. 12, 869.

Friedmannová L, Stanek K (2010) Cartographically augmented reality. Proceedings of the 3rd ISDE digital earth summit – digital earth in the service of society: sharing information, building knowledge. Nessebar, Bulgaria: ISDE.

Gilbert RL (2011) The prose (psychological research on synthetic environments) project: Conducting in-world psychological research on 3d virtual worlds. Journal of Virtual Worlds Research. 4, 1.

Goodchild MF (2010) Twenty years of progress: GIScience in 2010. Journal of Spatial Information Science. 1, 3.

Goodchild MF (2012) The future of Digital Earth. Annals of GIS. 18, 93.

Haklay M, Singleton A, Parker C.: 2008, Web mapping 2.0: The neogeography of the geoweb. Geography Compass. 2, 2011.

Hall GB, Chipeniuk R, Feick RD, Leahy MG, Deparday V (2010) Community-based production of geographic information using open source software and web 2.0. International Journal of Geographical Information Science. 24, 761.

Hirtle S, Maceachren A (1998) Cognitive models of dynamic geographic phenomena and their representations.
URL: http://www.ncgia.ucsb.edu/Publications/Varenius_Reports/Cognitive_Models.pdf.

Hu MY, Lin H, Chen B, Chen M, Che WT (2011) A virtual learning environment of the Chinese University of Hong Kong. International Journal of Digital Earth. 4, 171.

Huang B, Jiang B, Lin H (2001) An integration of GIS, virtual reality and the Internet for spatial data exploration. International Journal of Geographical Information Science. 15, 439.

Johnson A, Thompson EM, Coventry KR (2010) Human perception, virtual reality and the built environment. Proceedings of the 14th International Conference Information Visualisation (IV). The Institute of Electrical and Electronics Engineers (IEEE), California, USA, 604.

Kohler T, Matzler K, Fuller J (2009) Avatar-based innovation: Using virtual worlds for real-world innovation. Technovation. 29, 395.

Konecny M (2011) Review: Cartography: challenges and potential in the virtual geographic environments era. Annals of GIS. 17, 135.

Kraak MJ (2004) The role of the map in a Web-GIS environment. Journal of Geographical Systems. 6, 83.

Kreijns K, Kirschner PA, Jochems W, Van Buuren H (2007) Measuring perceived sociability of computer-supported collaborative learning environments. Computers & Education. 49, 176.

Lifton J, Paradiso JA (2012) Dual reality: Merging the real and virtual. http://dspace.mit.edu/openaccess-disseminate/1721.1/61991.

Lin H, Zhu Q (2005) Virtual geographic environments. In Zlatanova S, Prosperi D, (eds.) Large-scale 3d data integration: Challenges and opportunities. New York: Taylor & Francis. 211.

Lin H, Chen M, Lu G, Zhu Q, Gong J, You X, Wen Y, Xu B, Hu M (2013a) Virtual Geographic Environments (VGEs): A new generation of geographic analysis tools. Earth-Science Reviews. 126, 74.

Lin H, Chen M, Lv GN (2013b) Virtual geographic environment - a workspace for computer-aided geographic experiments. Annals of the Association of American Geographers. 3, 465.

Mania K, Badariah S, Coxon M, Watten P (2010) Cognitive transfer of spatial awareness states from immersive virtual environments to reality. ACM Transactions on Applied Perception. 9, 1.

Marchant JC (2010) Are you ready for life in world 3?. http://www.newscientist.com/blogs/culturelab/2010/08/are-you-ready-for-life-in-world-3.html. Retrieved 10 November 2012.

Meng L (2010) Changing world and its changing mapping strategies. http://www.szdcec.com/icgit2010/pdf/1018_1.pdf. Retrieved 05 October 2013.

Messinger PR, Stroulia E, Lyons K, Bone M, Niu RH (2009) Virtual worlds – past, present, and future: New directions in social computing. Decision Support Systems. 47, 204.

Miller HJ (2003) What about people in geographic information science? Computers, Environment and Urban Systems. 27, 447.

Montello DR, Waller D, Hegarty M, Richardson AE (2004) Spatial memory of real environments, virtual environments, and maps. In Allen GL, editors. Human Spatial Memory: Remembering Where. Mahwah: NJ, Lawrence Erlbaum Associates. 251.

NRC (2007) Linked social-spatial data: Promises and challenges. In Gutmann MP, Stern P, (eds.). Putting people on the map: Protecting confidentiality with linked social-spatial data. Washington: National Academy Press. 7.

Outram C, Ratti C, Biderman A (2010) The Copenhagen Wheel: An innovative electric bicycle system that harnesses the power of real-time information an crowd sourcing. http://cgt.columbia.edu/files/papers/Outram_Ratti_Biderman_EVER2010_Monaco.pdf. Retrieved 07 September 2013.

Pelachaud C, Martin JC, André E, Chollet G, Karpouzis K (2007) Spatial social behaviour in second life. Proceedings of the 7th International Conference on Intelligent Virtual Agents. Berlin/Heidelberg: Springer. 252.

Ragan E, Bowman D, Huber K (2012) Supporting cognitive processing with spatial information presentations in virtual environments. Virtual Reality. 1.

Ravitz G, Shyu ML, Powell MD (2010) Integrating multimedia semantic content analysis of YouTube videos with hurricane wind analysis for public situation awareness and outreach. International Journal of Software Engineering and Knowledge Engineering. 20, 155.

Resch B, Lipson S, Bers J, Ratti C, Mittleboeck M (2013) Integrated Urban Sensing: 2013, A Geosensor Network for Public Health Monitoring and Beyond. http://dspace.mit.edu/openaccess-disseminate/1721.1/64636. Retrieved 05 October 2013.

Scharl A (2007) Towards the geospatial web: Media platforms for managing geotagged knowledge repositories. In Scharl A, Tochtermann K, editors. The geospatial web - how geo-browsers, social software and the web 2.0 are shaping the network society. London: Springer. 3.

Shaw SL, Yu H (2009) A GIS-based time-geographic approach of studying individual activities and interactions in a hybrid physical–virtual space. Journal of Transport Geography. 17, 141.

Sheppard R, Cizek P (2009) The ethics of Google Earth: Crossing thresholds from spatial data to landscape visualization. Journal of Environmental Management. 90, 2102.

Sui D, Goodchild M. (2011) The convergence of GIS and social media: challenges for GIScience. International Journal of Geographical Information Science. 25, 1737.

Tuzun H, Yilmaz SM, Karakus T, Inal Y, Kizilkaya G (2009) The effects of computer games on primary school students' achievement and motivation in geography learning. Computers & Education. 52, 68.

Wilson D, Bertolotto M, Weakliam J (2010) Personalizing map content to improve task completion efficiency. International Journal of Geographical Information Science. 24, 7410.

Wood D, Fels J, Krygier J (2010) Rethinking the power of maps. New York: The Guilford Press.

Yang C, Raskin R, Goodchild M, Gahegan M (2009) Geospatial cyberinfrastructure: Past, present and future. Computers, Environment and Urban Systems. 34, 264.

Yu C, Peuquet DJ (2009) A GeoAgent-based framework for knowledge□oriented representation: Embracing social rules in GIS. International Journal of Geographical Information Science. 23, 923.

Cartographic Symbols Depicting Ecoregion Properties

Stella Whisler Todd[1]

Abstract The goal of this study was to develop a visual language about ecoregion properties which allowed for customisation at various spatial scales and delineations. The sub-objectives of this study were to: 1) determine a range of data values for the biophysical variables typical of global ecoregions, 2) develop a methodology to create multivariate symbols depicting ecoregion properties, and 3) demonstrate the symbol design with a set of Colorado ecoregions. Multivariate symbols were created to represent the complex array of biophysical properties associated with ecoregions, including temperature, precipitation, landforms, lithology, and representative vegetation. The symbol design accommodated more than 23 temperature and precipitation levels and multiple nested landform categories. Further research into the perception and cognition of sub-elements, as well as their application to a variety of ecoregion definitions will help to hone this prototypical symbol framework. The development of multivariate symbols is a concise tool to communicate complex interrelationships in a way that is not possible with single-variable symbols.

1 Introduction

A symbolic language can help to bridge verbal language barriers. Cartography utilises symbols to communicate geographic variation so that situations and events are recognisable and understandable regardless of our native language. Ecoregions are global entities of great importance to humanity and are natural geographical areas for managing natural resources and environmental quality. They are complex assemblages of unique biological, physical, and cultural properties. Symbols can convey properties within a concise format, viewable in maps of any dimension.

Cartographers depict regional diversity through the use of symbolic representations such as colours, dots (Slocum, 2009) or easily understood pictorial elements (Modley and Myers 1976, MacEachren 1994). The single thematic map remains

S. W. Todd
M.S., M.C.P, Ph.D., Metropolitan State University of Denver
Campus Box 22, P.O. Box 173362
Denver, CO 80217-3362, USA
e-mail: toddst@msudenver.edu

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_16,

the primary map type except for general reference maps containing many variables that are represented by markers or lines, with mutually exclusive areas. While most symbols depict specific things or events, some depict multivariate data as well as quantitative differences.

Multivariate cartographic symbols tend to be abstract representations of multiple factors (Dorling 2012, Grinstein et al. 1992, Nelson and Gilmartin 1996). Small multiple symbols depict comparisons between object values (Tufte 1990). Scenic visualisation of multivariate elements provides an intuitive context for displaying a variety of data, but the cognitive processes by which scenes are interpreted are not well understood (Hillery and Burton 2011). Displaying multiple factors within a single map poses unique challenges as it is difficult to depict more than one factor associated with a single area simultaneously. The lack of multivariate symbols could be due to symbol size or perception complexity issues, as well as lack of a rationale for combining variables in a single symbol. Viewing the association of many variables together is often the impetus for scientific hypotheses and investigation, however, as well as a means for understanding geographic differences.

Because each region occupies a unique geographic area a single symbol should be developed that uniquely identifies the area through the differences in the multiple sub-elements that constitute the region. Biophysical, land use and cultural factors vary between ecoregions. For example, temperature, humidity, terrain, lithology, and vegetation (or ecosystems) are commonly examined biophysical factors known to interrelate. Terrestrial ecoregions of the world are by definition multivariate and are therefore not well represented in traditional cartographic methods that focus on few variables. Each ecoregion represents a conglomerate of attributes that are internally similar, and distinct from other ecoregions. Multivariate ecoregion symbols were previously developed for Colorado (Todd 2012, Todd and Whalen 2011). No symbols depicting global variation in ecoregion properties have been developed where the biophysical sub-elements are distinguishable.

Classification methods for determining ecoregion boundaries may be systematic or intuitive (Jepson and Whittaker 2002). Manual interpretation of region boundaries is less explicit than systematic approaches but allows for local and/or holistic understanding (Chapman et al. 2006, Olson et al. 2001, Omernick 1987). Interpreting ecoregion boundaries can be as much of an art as a science, particularly when human land use is considered. The character of a region is always more robust than the variables we choose to measure or our ability to measure them. For this reason the manual interpretation of data sources as well as local knowledge may be justified.

In recent decades digital processing methods have commonly been used to construct and revise our concept of regions and ecosystems. Recent publications describe ecosystems of the conterminous US as well as African ecosystems at a 30m spatial resolution with input data from multiple digital data sources (Sayre et al. 2009, Sayre et al. 2013). Ecosystem modelling considers a number of biophysical properties as fundamental building blocks, such as ecological divisions, isobioclimate, lithology, landforms, topographic moisture potential, and vegetation cover.

Finely-resolved ecosystem datasets afford a good opportunity to quantify the biophysical parameters of ecoregions in a meaningful way. They allow for the addition of detailed information about biophysical distribution patterns within established ecoregions (Olson et al. 2001, Omernick 1987, Bailey 1996). Landform categories include high mountains, low mountains, foothills, escarpments, irregular plains, smooth plains, flat plains, and drainages. Terrain relief can be estimated from these categories. Isobioclimate categories are derived on a global scale as well, with corresponding precipitation and temperature values. The most dynamic ecoregion category is ecosystem type as hundreds of ecosystems have been identified (Sayre et al. 2009, Sayre et al. 2013). Ecosystems and landforms are categorical but metrics for their proportional composition within ecoregions may be depicted within nested or multiple sub-elements.

In order to better understand and recognise the unique characteristics of ecoregions, a standardised set of cartographic symbols representing the inherent variability of biophysical factors would be beneficial. Symbols are an important aspect of cartographic images. Our human understanding of an image transcends the words that describe it. Because ecoregions are globally distributed it makes sense to develop a standardised symbolic language to communicate their importance, uniqueness, and dynamics to scientists and non-scientists internationally. Standardisation does not preclude flexibility in adapting symbols to different classification methods but encourages commonality in symbol interpretation for specific sub-elements.

2 Objectives

The overall goal of this study was to standardise ecoregion representation to develop a visual language about ecoregions while allowing for customisation for various spatial scales and applications. The sub-objectives of this study were to: 1) determine the range of data values for biophysical variables typical of global ecoregions, 2) develop a methodology to create multivariate symbols of global ecoregions depicting their properties, and 3) demonstrate the symbol design with a set of Colorado ecoregions.

3 Datasets

Primary GIS datasets were acquired in order to estimate a range of data values for climate and terrain variables and identify other biophysical variables important in ecoregion delineation (Table 16.1). Ecosystems and their component inputs were obtained from USGS (Sayre et al. 2009). Ecoregion boundaries were obtained from a Level III and Level IV ecoregions dataset from the US EPA as described in Chapman et al. (2006).

Table 16.1 Project datasets containing biophysical variables, ecoregion boundaries, and land cover information.

Dataset	Variable	Source	Description
co_eco_14.shp	Ecoregion boundaries	http://www.epa.gov/wed/pages/ecoregions/co_eco.htm	Ecoregions of Colorado
ecosystems_ibio.img	Isobioclimates	http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml	Isobioclimate layer Colo eco bnd
ecosystems_landf.img	Landforms	http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml	Landform layer Colo eco bnd
ecosystems_lith.img	Lithology	http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml	Lithology layer Colo eco bnd
Final_labeled_prod4_mmu9pxl_30m_dd83_w_subsec.img	Ecosystems	http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml	Entire US ecosystems from NatureServe
US_ecosys_lookup_tables.xls	Metadata	http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml	Links to ecosystem and component codes
Landcovi2001.tiff	National Atlas Land Cover	http://nationalatlas.gov/atlasftp.html?openChapters=chpbio#chpbio	Land Cover Map at 200m scale from 2001

4 Methods

The following criteria were considered in symbol design: 1) the symbol should be small enough to be used as a map symbol and/or within a small application interfaces (3 cm x 3 cm), 2) The sub-elements should be individually distinguishable, and 3) the sub-elements should be modular, 4) the symbol sub-elements and levels should be identifiable and distinguishable from one another in greyscale, 5) the symbol sub-elements should be able to depict variation in climate, landform, and lithology worldwide, 6) the symbol prototype should be useful at a variety of spatial scales, and 7) the symbol design should be able to accommodate differences in ecoregion delineations among agencies.

5 Creation of Multivariate Symbols

In order to develop a symbolic prototype of ecoregions, the delineations of Omernick, as described by Chapman et al. (2006), were initially cross-tabulated with the ecosystems and other datasets of Sayre et al. (2009) in order to grasp the data ranges within Colorado. The design was then modified to depict global data values.

Isobioclimate categories were translated into quantitative variables using the Ombrothermic Index (Io) values representing precipitation and positive temperature (Tp) representing heat load from Rivas-Martinez et al. (2009). The minimum Io value within the data range of each main type was used to populate the field in each case, within 0.1 decimal places. The minimum value was chosen because the data range for the most humid class (Io > 24) was unknown, whereas the minimum was close to zero. Ombrotype values represent variation in precipitation within ecoregions. The minimum Io value within the data range of each main type was used to populate the field as some value ranges were classified as > x. The thermotypes represent a combination of two variables: positive temperature (Tp) and Thermicity Index (Itc). The minimum Tp value within either the lower thermotype or the upper thermotypes represented the thermotype. The minimum value was selected because the upper limits were not specified. The Tp value was selected to represent thermotype values for symbol development purposes.

6 Colorado Multivariate Symbol Development

Symbols were constructed for ecoregions across the mountainous centre of the state using an east-west transect to illustrate variation in ecoregion properties (Figure 16.1). Selected Level IV ecoregions included portions of the Southern Rockies and High Plains (Level III) ecoregions described by Chapman et al. (2006). The Southern Rockies is a diverse geographic area with high elevations and steep terrain. A range of cover types is present, from grasslands to forests, with dramatic changes in the visual landscape along elevation gradients. Within the Southern Rockies geographic area selected Level IV ecoregions include the Alpine Zone (21a), Crystalline Subalpine Forests (21b), Crystalline Mid-elevation Forests (21c), Foothill Shrublands (21d), Sedimentary Subalpine Forests (21e), and Sedimentary Mid-elevation Forests (21f). The High Plains are extensively managed areas of grasslands, rangelands, pasturelands, and croplands. Level IV ecoregions within them include the Rolling Sand Plains (25b), Flat to Rolling Plains (25d), and Front Range Fans (25l).

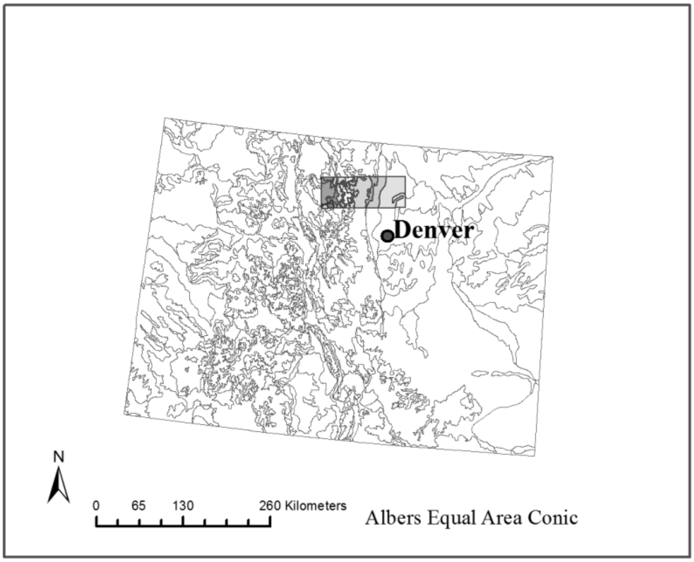


Fig. 16.1 Locations of Level IV ecoregion transects in Colorado

Symbol sub-elements were designed as follows:

1. Each cloud incremented the Ombrothermic Index (Io) value by 1, with the global data range from 1 to 24. For Io values less than one, up to 10 dots construct a single cloud. For example, an Io value of 2 was represented by 2 clouds and an Io value of 0.2 was represented by two dots within a single cloud outline.
2. Each sun ray incremented the Positive Temperature (Tp) by 100 starting at Tp = 100, with the global data range from 1 to 2400 (24rays). A Tp of 2700 was represented with an additional inner circle. For example a Tp of 400 had 4 rays.
3. The ten landform (terrain relief) categories were depicted as individual polygons in such a way that all categories would be visible even if nested. This allowed for more than one category to be depicted simulta-

neously. The terrain categories occupying at least 10% of the ecoregion were all depicted within a single symbol. Landform height on the symbol (mm) was designed proportionally to terrain relief (m) (R^2=0.97).

4. The most prevalent lithology class was symbolised as a unique texture.
5. Up to three representative species were selected based on ecosystem and/or land use descriptions. For example three representative plant species were selected based on the criteria below:
    a. Plant species descriptions within the most prevalent ecosystem within each ecoregion (Sayre et al. 2009).
    b. Ecoregion description of dominant vegetation and identifiable species (Chapman et al. 2006).
    c. The predominant land use or land cover class within each ecoregion using the 2001 National Atlas Land Use Land Cover Map.
    d. Pictorial elements of representative vegetation were created from simplified line drawings of vegetation photographs.

7 Results

It was possible to create multivariate symbols of ecoregions depicting their properties by simplifying and scaling biophysical sub-elements (Figure 16.2). Multivariate symbols were created to represent the complex array of the biophysical properties of terrestrial ecoregions including temperature, precipitation, landforms, lithology, and representative vegetation (for ecosystems). The range of data values for climate and terrain variables typical for global ecoregions was determined. These were demonstrated for Colorado Omernick Level IV ecoregions. Areas of the Colorado Southern Rockies and High Plains were represented in the sample transect. More than one landform was allowed within a single symbol in this prototype, as each landform was visible even when stacked. The landform symbols clearly showed a change in landscape features from the plains to the mountains. The inclusion of more than one landform depicted the variation in landform structure within each ecoregion. This resulted in the same visualisation for the alpine zone (21a) and crystalline subalpine forest (21b) landforms even though the proportion of these landforms varied between the two ecoregions (Table 16.2). Three dominant vegetation species were visually distinct and were small enough to obscure only portions of the landform sub-elements. The temperature sub-elements were easily seen and the rays countable, with a single temperature metric representing the region. Precipitation sub-elements were likewise easily discernable and simple to count.

Table 16.2 Biophysical properties of Colorado Level IV ecoregions, as described in Chapman et al. (2006), using digital data from Sayre et al. (2009), including landform category proportions, predominant lithology category, isobioclimate values (positive temperature (Tp) and ombrothermic values (Io)), and reference vegetation from ecoregion descriptions, ecosystem descriptions, and land use land cover maps (nationalatlas.gov).

	21a	21b	21c	21d	21e	21f	25b	25d	25l
Flat Plains	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.63	0.22
Smooth Plains	0.00	0.00	0.00	0.01	0.00	0.01	0.45	0.30	0.33
Irregular Plains	0.01	0.01	0.04	0.14	0.04	0.09	0.10	0.04	0.30
Escarpments	0.01	0.01	0.01	0.03	0.02	0.03	0.00	0.00	0.01
Hills	0.00	0.01	0.04	0.05	0.01	0.03	0.00	0.00	0.02
Foothills	0.02	0.10	0.17	0.23	0.13	0.20	0.00	0.00	0.03
Low Mountain	0.47	0.57	0.49	0.34	0.55	0.46	0.00	0.00	0.02
High Mountains	0.36	0.12	0.04	0.01	0.09	0.02	0.00	0.00	0.00
Drainage Channels	0.12	0.18	0.20	0.18	0.15	0.17	0.02	0.04	0.08
Lithology Class	SRM	SRM	SRM	NCRM	NCRM	NCRM	SLS	NCRM	NCRM
Average (Tp) Minimum	736	630	626	639	655	607	815	883	800
Average (Io) Minimum	8.44	5.89	4.47	3.18	5.69	4.05	2.23	2.11	2.33
Veg (Ecosystem)	Picea engelmannii (Krumholz)	Abies lasiocarpa	Pseudotsuga Carrière	Grassland	Abies lasiocarpa	Pseudotsuga Carrière	Andropogon gerardii	Shortgrass	Shortgrass
Veg (Ecoregion)	Grassland	Grassland	Pinus contorta	Artemisia	Grassland	Quercus gambelii	Artemisia filifolia	Tallgrass	Tallgrass
Veg (National Atlas)	Grasslands/herbaceous	Evergreen Forest	Evergreen Forest	Grasslands/herbaceous	Evergreen Forest	Evergreen Forest	Grasslands/herbaceous	Grasslands/herbaceous	Grasslands/herbaceous

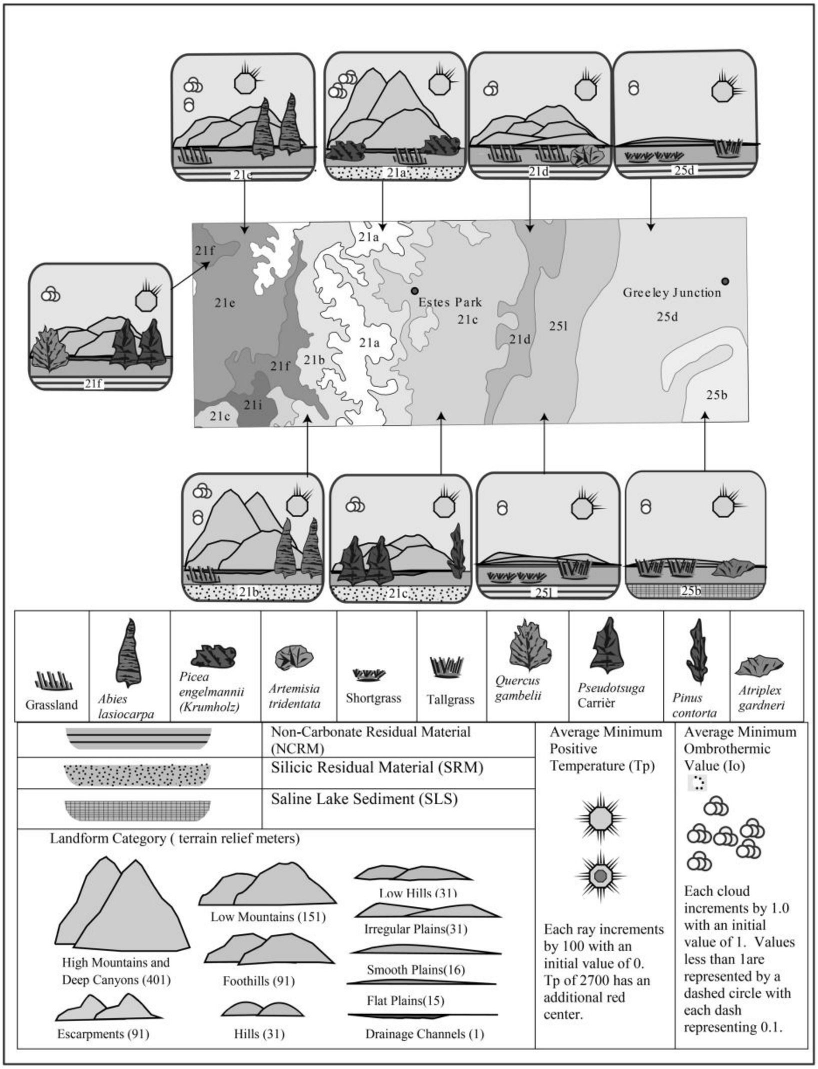


Fig. 16.2 Colorado ecoregions and associated symbols depicting biophysical properties

8 Conclusions

Developing a common visual symbolic language to represent ecoregions helps to present the biophysical complexity of these regions in a simplified form. The development of multivariate symbols communicates complex interrelationships not possible with individual symbols for each variable. The abstraction necessary to depict variation at a global scale reduces sub-element distinctions for a small geographic area such as Colorado; the temperature range for Colorado, for example, is much smaller than the global temperature range. Representative vegetation is the most subjective sub-element since vegetation structure is complex and the geographical scale of variation for ecosystems is smaller than for the other biophysical properties.

Biophysical variables such as climate, landforms, geologic or soil structure, and ecosystems will vary among ecoregions and will vary with the process by which ecoregions are defined, and therefore the symbolic depiction of these properties will also need to vary to some extent. The development of standardised symbolic depictions of biophysical properties will be instrumental for conveying ecoregion differences globally. Understanding the unique differences in biophysical variables allows decision-makers to consider the impact of social and political actions on these areas. Visualising these differences through the development of cartographic symbols communicates the uniqueness and importance of ecoregions in land use decisions.

9 Discussion

The symbol development process for this project assumed that ecoregion structure was an important consideration in delineating and representing ecoregions. The importance of these structural elements varies between agencies that develop ecoregion boundaries. Even if a common biophysical structure is considered, the variables may be altered to accommodate a specific spatial scale or classification hierarchy. For example, rather than lithology, soil type might be a more instructive variable for a small region.

Another consideration is how to depict a range of data values, or the proportion of identified types for one attribute within a symbol. Some ecoregions may consist of ten ecosystems or plant communities, for example, while others only consist of one or two types. A proportional cut-off of 10% may not be the ideal threshold for categorical variables in order to be represented as a symbol sub-element. The decision about whether to represent presence, or a minimum, maximum, or mean value, or the variation in data values, will alter the utility of the resultant symbol. For example, the mean terrain relief may be included on the symbol as a line in addition to pictorial depictions of landforms.

Symbolising ecosystems, land use or cultural categories is the most difficult and subjective aspect of developing a visual language for ecoregions. Each ecosystem contains myriad species with more than one dominant species possible. If only one species is chosen to represent a single ecosystem the same species may be chosen to represent another ecosystem as well, so there is potential confusion in legends. Simplifying the numerous vegetation species to a small yet distinguishable sub-element is an arduous task, but it is not impossible. Land uses and cultural importance is difficult to ascertain and favours intuitive judgment. Additional sub-elements may be warranted for ecoregions based on biotic and human factors, such as endemic species, endangered species, or economic activities. This would require a larger symbol size or other modifications to the prototype presented in this study.

Abstraction and simplification is a third consideration to evaluate. Although simplification may convey the correct data categories, the symbol may not be similar enough to the actual landscape to be meaningful to the public. A photograph or artistic drawing of the landscape might be more helpful in communicating the uniqueness and value of a region than an abstracted landscape. Further research into perception and the cognition of sub-elements, as well as its application to varied ecoregion definitions, will help to hone this prototypical symbol framework.

Developing one or more standardised visual systems of symbols for ecoregions worldwide is an endeavour worthy of consideration. Symbols have the potential to stimulate conversation and understanding about ecoregions. They can be helpful in determining and communicating the factors which are important in their delineation, as well as the differences between regions. Changes in ecoregion structure due to land use and other human activities, as well as climate alteration, can be visualised in order to raise awareness of their importance and dynamics.

References

Bailey RG (1996) Ecosystem geography. Springer, New York

Chapman SS, Griffith GE, Omernik JM, Price AB, Freeouf J, Schrupp KL (2006) Ecoregions of Colorado (color poster with map, descriptive text, summary tables, and photographs): U.S. Geological Survey (map scale 1: 1,200,000), Reston

Dorling, D (2012) The visualization of spatial social structure. Wiley, West Sussex

Grinstein G, Sieg JCJ, Smith S, Williams MG (1992) Visualization for knowledge discovery. International Journal of Intelligent Systems 7: 637-648

Hillery BA, Burton RP (2011) Multivariate data visualization via outdoor scenes. In:Wong PC, Park J, Hao MC, et al (eds) Visualization and Data Analysis. SPIE-IS&T vol 7868. doi: 10.1117/12.872661

Jepson P, Whittaker RJ (2002) Ecoregions in context: a critique with special reference to Indonesia, Conservation Biology 16(1): 42-57

MacEachren AM (1994) Some truth with maps: A primer on symbolization and design. Association of Geographers, Washington DC

Modley R, Myers WR (1976) Handbook of pictorial symbols. Dover Publications, New York

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D’amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR, (2001) Terrestrial ecoregions of the world: A new map of life on Earth. Bioscience 51(11): 933-938

Omernick JM (1987) Ecoregions of the conterminous United States: Annals of the Association of Geographers 77(1): 118-125

Nelson ES, Gilmartin P (1996) An evaluation of multivariate quantitative point symbols for maps. In: Wood CH, Keller CP (ed.) Cartographic Design: Theoretical and Practical Perspectives. Wiley, Chichester, pp191-210

Rivas-Martinez S, Rivas y Sáenz S (2009) Synoptical Worldwide Bioclimate Classification System (summarized table): Centro de InvestigacionesFitosociologicas. http://www.globalbioclimatics.org/book/namerica2/table.htm. Accessed 15 Jan 2014

Sayre R, Comer P, Hak J er al (2013) A new map of standardized terrestrial ecosystems of Africa. Association of American Geographers, Washington DC

Sayre R, Comer P, Warner H, Cress J (2009) A new map of standardized terrestrial ecosystems of the conterminous United States. U.S. Geological Survey Professional Paper 1768

Slocum TA, McMaster RB, Kessler FC, Howard HH (2009) Thematic cartography and geovisualization, 3rd ed. Prentice Hall, New Jersey

Todd SW (2012) Visualizing ecoregions. Poster presented at the National American Cartographic Information Society Annual Meeting. October 17-20, Portland 2012

Todd SW, Whalen B (2011) Multivariate symbology of Colorado ecoregions. Poster with abstract presented at the National American Cartographic Information Society Annual Meeting, October 12-14, Madison

Tufte ER (1990) Envisioning information. Graphics Press, Cheshire, CT

Preparing Simplified 3D Scenes of Multiple LODs of Buildings in Urban Areas Based on a Raster Approach and Information Theory

Alexey Noskov and Yerach Doytsher[1]

Abstract We have developed a method for the simplification of the footprints of 2D buildings based on a rasterisation process. The rasterisation is processed within quarters and the urban area is subdivided into quarters based on natural contours such as roads and water objects and not on straight geometric lines (the common subdivision approach to orthogonal tiles). Quarters were organised into a hierarchical model, according to the gaps between the quarters and the stages of the clustering process, using Kohonen's self-organising maps. Each degree of simplification (generalisation) corresponds to some level of hierarchy. Information theory was used to estimate the amount of the 2D generalisation of building footprints. Simplified building footprints were extruded for the compilation of a 3D urban perspective from multiple levels of detail (LODs) The entropies of 3D scenes for each quarter of the hierarchy and each LOD were compared in order to define the level of detail to be used in the final 3D scene.

1 Introduction

The prevalence of computer devices (especially smartphones and tablet computers) and the latest achievements in software, enable us to use 3D maps almost everywhere. The two most common problems to arise in any discipline are that: (1) huge computer resources are required for drawing 3D models based on the original, non-simplified models, and (2) 3D models based on the original non-simplified objects are very detailed and often appear unreadable and overly complex. Some method for the simplification of models has to be applied to resolve both problems. The simplification of urban area maps is an extremely complex topic, mainly because of the variety and complexity of building models. The main

A. Noskov and Y. Doytsher
Mapping and Geo-Information Engineering, Technion – Israel Institute of Technology
Haifa, Israel
e-mail: mail@alexkarta.com

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_17,

types of objects in a 3D city model are buildings, and so this chapter focuses on the simplification of buildings. There are two different tasks in the building simplification process: (1) simplification of a single building, and (2) generalisation of groups of buildings. “Simplification of a single building” has been widely researched as a subject, and we can describe several different approaches to generalisation, all of them valid. In contrast, “generalisation of a group of buildings” has been treated, so far, only on a very limited level. There are several very similar approaches, mainly based on the Delaunay Triangulation (DT) (e.g., Li et al. (2004) and Xie et al. (2012)). We propose, in concept, another approach which enables the simplification of buildings and groups of buildings in one holistic process. The approach is based on rasterisation and vectorisation operations, which are carried out by sub-dividing the urban neighbourhood into quarters. Initial results for this approach were presented by the authors in Noskov and Doytsher (2013a). In this chapter, we suggest the use of information theory to estimate the quality of the simplification of 2D building footprints. Information theory is also used to define the optimal levels of detail (LOD) of the buildings in each city quarter depicted in the final 3D scene. This chapter is structured as follows: related work is considered in Section Two, the algorithms of raw quarter calculations and building quarter hierarchy are presented in Section Three, the raster-based algorithms for generalising a group of buildings is considered in Section Four, evaluation of the generalisation results are presented in Section Five, and the new information-based content analysis approach to compiling 3D scenes is presented in Section Six, while concluding results are presented in Section Seven.

2 Related Work

The holistic approach to the 3D generalisation of buildings was described in Xie et al. (2012). The main idea supposes that, within a threshold (distance from a viewpoint), we will generate objects which contain the results of the simplification of single buildings, whereas outside the threshold we will generate objects containing the results of groupings of buildings and their simplification as a single building. An approach to “converting 3D generalisation tasks into 2D issues via buildings' footprints” was described in He et al. (2012).

The generalisation of 3D building data approach (Forberg 2007), based on scale-space theory from image analysis, allows the simplification of all orthogonal building structures in one single process. Another approach (Thiemann 2002) considers buildings in terms of constructive solid geometry (CSG). In Xie et al. (2012) an approach was proposed which realised 3D single building simplification in five consecutive steps: building footprint correction, special structure removal, roof simplification, oblique facade rectification and facade shifting. A very interesting approach was proposed in Kada (2002). In this approach, geometric simplification was realised by remodelling the object by means of a process similar to half-space modelling. Approximating planes were determined from the polygonal faces of the original model, which were then used as space-dividing primitives to create facade and roof structures of simpler shapes.

The second aspect of the 3D generalisation of an urban environment is the generalisation of groups of buildings. 3D generalisation of groups of buildings is mentioned in several publications (e.g. Glander and Döllner (2009), Guercke et al. (2011), He et al. (2012), and Trapp et al. (2008)). These papers describe different approaches to 3D grouping and group generalisation: the grouping of building models (using the infrastructure network) and replacing them with cell blocks, while preserving local landmarks (Glander and Döllner 2009); "express[ing] different aspects of the aggregation of building models in the form of Mixed Integer Programming problems" (Guercke et al. 2011); and, the grouping of building models "with a minor height difference and the other with a major height difference" (He et al. 2012).

In spite of the large number of publications and developing projects, we can identify a very important shortcoming in almost all 3D city maps (or screenshots). There are usually only two ways of displaying large cities: depicting only the buildings nearest to the viewpoint (where all the other buildings are not displayed and the area is depicted as a plain map), or displaying all the 3D building models (which usually involves a lengthy processing time and heavy computer and internet traffic resources, and furthermore, causes distant parts of the city to be presented as a very dense and an unreadable 3D view). As mentioned above, there are a large number of publications on generating and using the LODs of single buildings, while there are only a very limited number of approaches to creating LODs of group of buildings in order to solve the problem described above.

There are also related links in studies used to calculate the entropy of simplified layers and 3D scenes. The 'entropy' is the main measure used to estimate the quality of simplification. As noted by Shannon (1948), 'entropy' is a quantitative measure of information content. Sukhov (1970) implemented a quantitative measure of the information content of a map for the first time. Neumann (1994) measured the topological information of a contour map. The effectiveness of map design and the information content of maps were considered in Knopfli (1983) and Bjorke (1996). To estimate the quality of a map and the compiled 3D scene, we used three measures: coordinate digit density (CDD) function (Battersby and Keith 2003, and Clarke et al. 2001), entropy of Voronoi regions, and entropy of Voronoi neighbours (Li and Huang 2002); for more details see Section Five.

3 Calculation of Raw Quarters and Quarter Hierarchy

Finding a practical method of simplification is a very important issue in generalisation. One of the more common problems occurs when buildings are joined through obstacles such as wide roads or rivers. In this case, buildings must not be joined to each other, and those buildings from the two sides of the obstacle should be merged with other, more distant objects, which are, however, located on the correct side of the obstacle. To resolve this problem, we decided to split the urban space into quarters which are divided by the main, significant objects. These objects cannot be involved in the generalisation itself.

To calculate the quarters, we decided to use the slope of the terrain, water objects and roads. For implementing and testing our approach, geodata for the city of Trento, Italy was used. The buildings (with individual heights), water objects and roads were extracted from the landuse map of Trento. The landuse map and land relief (DEM) were downloaded from the website of Trento Municipality, Italy. It was found that buildings were positioned only in areas with a slope of less than 30 degrees. Areas of the terrain with slopes greater than 30 degrees were thus excluded. We also used roads and water objects for defining the quarters. These objects are polygons extracted from the landuse map. All three classes – slopes, roads and water objects - were merged into one raster map with 1 meter resolution (which has been found to be adequate for small-scale urban generalisation). All pixels of the merged objects were given the value “1”; an empty space on the raster map was given a “0” value. Each group of pixels with value “1” was expanded by adding one pixel (1 meter). In the next step, the values of pixels were inverted (“1” to “Nil” and vice versa), resulting in many pixel areas with the value “1” which were split by “Nil” pixels. The final raw quarter map was prepared by vectorisation of the inverted raster map. Fig. 17.1 illustrates the process of preparing quarters.

In the initial phase of the research we prepared raw quarters. We could not use the raw quarters for a high level of generalisation, as raw quarters would be too small for large generalised buildings. To overcome this limitation, a new, flexible hierarchical approach of subdividing the urban area into variable quarters, was developed. The raw quarters were placed on the lowest level of the hierarchy; on the highest level, the whole area of the city was defined as one large quarter. A special approach to developing the quarter generalisation (or quarters merging) enabled this hierarchy, where the size and content of the quarters are correlated with the level of the 3D generalisation, and this, in turn, was related to the distances from the viewpoint when compiling the 3D urban scene. Each level of the hierarchical tree of quarters had some level of quarter generalisation. The hierarchy was based on buffer operations. We widened the quarters with a buffer; thus, adjacent quarters were merged into one object, while other objects only changed their geometry (the outer boundary of quarters were simplified; some small inner elements, such as holes and dead-end roads, were filled). We then decreased the quarter by a buffer of the same size. If the width of the gap of merged quarters was smaller than the buffer width, the objects remained as a merged polygon; otherwise the polygon was split back into separate objects differing from the original objects due to their simplified geometry.

The first step of quarter generalisation was calculating the attributes for each quarter. These calculated multiple attributes (area, compactness, perimeter, fractal dimension of building objects and quarter areas, etc.) were used for the classification of the quarters. We decided to use Kohonen's Self-Organising Map approach to classify quarters. The number of clusters was defined for all levels in the hierarchy of the quarters to perform classification. There are several techniques to automatically define the numbers of clusters (e.g., a gap statistic approach, an information theoretic approach, etc.). At this time, we focused on a manual definition of the number of classes. An initial manual analysis of a series of maps based on quarter attribute visualisation was carried out. For more information about the described classification method see Noskov and Doytsher (2013b).

Fig. 17.1 A) Shaded landuse map of Trento and buildings; B) Non-nil pixel groups which split the city space into quarters (inverting pixel map); C) Final map of raw quarters (defining quarter areas having unique values)

As described above, adding a weighted buffer to the quarters has been suggested in order to take the quarter classes into account (thus differentiating between quarters of the same class and quarters of different classes), aiming to merge neighbouring quarters. To avoid vector artefact and topology problems, data was converted to raster, and buffering operations were executed in a raster environment. We used the raster resolution as the base width of the buffer. Not only does the buffering phase provide the possibility of merging quarters, this operation also helps to fill holes and dead-end roads in polygons, and to eliminate small elements of quarter boundaries. It should be noted that converting vectors to raster can also work as a generalisation operation. Generally speaking, this phase in the research, which was based on vector to raster and raster to vector operations, as well as region growing and buffer implementations on the one hand, and on the quarters' attributes on the other hand, enables us to generalise the quarters and lets us move up or down in the hierarchical level of the quarter subdivision. Quarters of the same class were merged faster than quarters of different classes. This was achieved by putting quarters of the same class into isolated sub-environments (temporal layers) and using different widths of buffers (see result in Fig. 17.2).

This suggested approach allows quarter generalisation based on buffering operations, while taking into account quarter classes. Using this method builds a hierarchical tree of quarters (in the current sample of Trento, from the raw source of 2679 small quarters up to a single huge quarter). We performed the generalisation of quarters starting from a buffer width of 1 meter and increasing it by increments of 0.2 meters, until the buffer width reached 2.6 meters. It was decided to start quarter generalisation from 1 meter because this is the resolution used to generate the raw quarters; and set the upper limit at 2.6 meters because a higher buffer width generates oversized quarters.

We decided to use 8 degrees of building generalisation based on rasterisation processes with resolutions of 10, 15, 20, 25, 30, 40, 50 and 60 meters (resolutions that correspond to degrees of generalisation). A graph of the varying number of quarters and the size of a maximal quarter was used to define which levels of hierarchy could be used for further processing. In addition, the original vector map of buildings was converted to raster maps with different resolutions (10, 15, 20, 25, 30, 40, 50 and 60 meters). These raster maps, overlaid with the generalised quarter maps, were used to estimate which resolution of buildings generalisation should be used with each generalised quarter map. Finally we decided to use a scheme of correspondence between buffer sizes used to calculate quarter levels and pixel sizes used for generalisation (buffer size – raster resolutions) as follows: 0.0 (raw quarters) – 10 and 15, 1.0 - 20 and 25, 1.2 – 30, 1.4 - 40, 1.6 – 50, 1.8 - 60 .

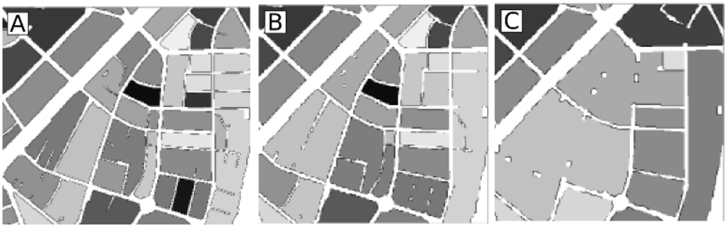


Fig. 17.2 Quarter buffering generalisation (buffer width): A) Original raw quarters; B) 1.4 meters; C) 1.8 meters

4 Simplification of Buildings

The fact that in urban areas, most (if not all) buildings have orthogonal sides, is the basis of our raster-based generalisation approach. In adjacent areas (quarters in our case), buildings would usually be spatially oriented in the same direction. The generalisation process therefore consists of defining the typical azimuth of a building's sides for each quarter. Once a typical azimuth is known, by applying the rasterisation process in this direction, the staircase-type appearance of lines, or legs of closed polygons, which is very common in the rasterisation processes, can be eliminated. A non-rotated rasterisation (parallel to the grid axes), while the buildings are positioned in another orientation, will result in the staircase-type appearance of the bordering lines of the buildings and in too many unnecessary vertices, which will prevent us from achieving a smooth geometry of the generalised objects.

Fig. 17.3 The Generalisation process of buildings in a quarter: A) Original buildings; B) Rotated quarter and the generalised 10 meter rasterised buildings in red; C) Final result

Within each quarter, the azimuths of all the building's sides were computed. For each building in the quarter, the longest side and its azimuth were identified, then all the azimuths of the other sides were rotated by 90 degrees (clockwise) again and again; and the rotated azimuths (and their lengths) were put into one list. The list was sorted by lengths, and then lengths with the same azimuths (up to a predefined threshold) were averaged. A threshold of 1 degree when looking for adjacent building side azimuths was found to give satisfactory results. A weighted average of the azimuths of the longest lengths of all the buildings within a quarter was used to define the general orientation of all the buildings of the quarter.

In order to significantly reduce the number of vertices of the generalised building, and achieve a more realistic appearance of these simplified objects, rasterisation should be carried out in the spatial orientation of the buildings. A rasterisation which is spatially oriented parallel to the grid axes will define buildings which are not oriented parallel to the grid axes in a staircase-type appearance of the buildings' sides. Accordingly, all the buildings within a quarter were rotated counter-clockwise at the angle of their general orientation to all the buildings of the quarter. The rotated buildings were then rasterised using a specific pixel size resolution (as explained in the next section). Each pixel with more than half its area covered by the original buildings was given the value "1"; otherwise, it was given the value "Nil". Figure 17.3 shows the results of this stage.

The level of generalisation is a function of the pixel size rasterisation process - the greater the pixel size, the greater the degree of generalisation. Accordingly, each quarter was generalised at several levels of rasterisation, resulting in several layers of different levels (level-of-detail) of generalised buildings for each quarter. Based on the original data from Trento, and according to our analyses, we found that using pixel size resolutions of 10, 15, 20, 25, 30, 40, 50 and 60 meters produces satisfactory results in the continuous and consecutive appearance of the level-of-detail of the generalised buildings. Generalised buildings were stored in separate layers; the identifiers of these layers contained the resolution of the building generalisation and the number of the level in the quarter hierarchy (or actually, the buffer width).

5 Evaluation of Quality of 2D Footprints Simplification

Eight LODs of buildings were calculated for the city area. The important issue is the evaluation of the results of simplification. The solution to the problem can be found by using information theory. Our attention concentrated on three ways of calculating the entropy of maps: (1) coordinate digit density (CDD) function, (2) entropy of Voronoi regions and (3) Voronoi neighbours. Entropy is a measure of the uncertainty of information content of a map.

The CDD method of calculating the entropy of map was described in Battersby and Keith (2003) and Clarke et al. (2001). The main idea is based on the calculation of the probability of digits in coordinate values of objects. The approach can be expressed with four equations:

$$P(d_n) = \frac{1}{Sn}\ (1);\ D_n = O(d_n) - P(d_n)\ (2);\ H(n) = \sum_{1}^{s} |D_n(s)|\ (3);\ I = \sum_{0}^{n} H(n)\ (4);$$

In Eq. 1 $P(d_n)$ is the probability of an individual character d in digit place n, Sn in a number of possible states. In Eq. 2 D is the overall variation of digit n. $O(d_n)$ is the probability of a digit for all points, $P(d_n)$ is the common probability of d in a considered numerical system. For a standard decimal numerical system, $P(d_n)$ is equal to 0.1. $H(n)$ (Eq. 3) is the measure of information potential and I is total information content (entropy) in a geographic dataset. The entropy of digits in place number 3, 4, 5 from X coordinates 66**445**6, 66**422**3, and 66**440**3 of 3 random centroids of simplified buildings could be calculated as follows:

Digit place #3: D(j)= 0/3-0.1=0.1, where j is 0-3,5-9; D(4)= 3/3-0.1=0.9; H(4,4,4) = 0.9 + 9*0.1=1.8

Digit place #4: D(j)= 0/3-0.1=0.1, where j is 0-3,5-9; D(2)= 1/3-0.1=0.23; D(4)=2/3-0.1=0.57; H(4,2,4)=1.6

Digit place #5: D(j)= 0/3-0.1=0.1, where j is 1,3,4,6-9; D(i)=1/3-0.1=0.23, where i is 0,2,5; H(5,2,0)= 1.4

I= 4.8

Another method involves calculating entropy using Voronoi polygons, reported in Li and Huang (2002). The entropy of Voronoi polygons could be calculated using the areas of polygons and the number of neighbours. The entropy of Voronoi regions is calculated as follows:

$$H = -\sum_{1}^{n} \frac{Si}{S} (\ln Si - \ln S) \qquad (5);$$

Si is the area of individual polygons, S is the whole area, n is the number of polygons. The number of polygons affects this measure; thus final entropy could be normalised:

$$H_N = \frac{H(M)}{\log_2 n} \qquad (6);$$

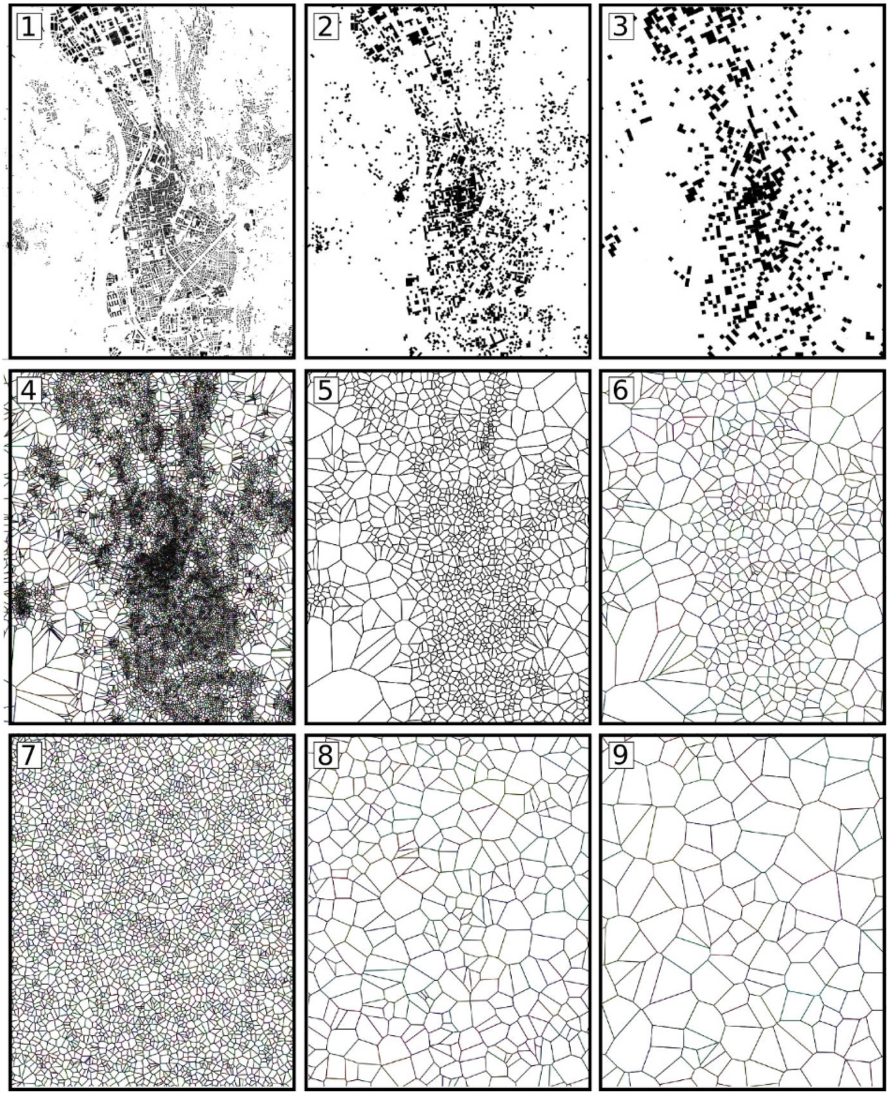


Fig. 17.4 Building layers and Voronoi polygon maps: 1-3) building maps (left-to right: original buildings, generalised by 30m pixel buildings and generalised by 60m pixel buildings); 4-6) Voronoi polygon maps based on points derived from centroids of polygons of the upper building map; 7-9) Voronoi polygon maps based on pseudo-random points, number of points equal to the number of centroids of the upper building map's polygons

To calculate entropy of Voronoi neighbours, the centroids of polygons are connected by Delaunay triangulation, and then the number of neighbours can be easily calculated for each centroid. The entropy is calculated as follows:

$$H = \sum_{j=1}^{Mj} \frac{n_j}{N_j} \ln\left(\frac{n_j}{N_j}\right) \qquad (7);$$

It was found that Voronoi polygons based on pseudo-random point maps could be used as a basic sample, which can be used to evaluate the results of generalisation. A pseudo-random generator from C standard library was used to calculate point coordinates. Two types of random maps were prepared for each level of generalisation and original building layer. Voronoi polygon maps were based on pseudo-random centroids, where the number of polygons is equal to the number of polygons of a correspondent building map. Point maps are where the number of points is equal to the number of vertices of the polygon boundaries of a correspondent map of buildings. The first type was used to evaluate the generalisation through entropy of the Voronoi polygon method. The second type was used to evaluate generalisation through the CDD function method. Figure 17.4 presents comparisons of pseudo-random Voronoi maps and generalised buildings layers' Voronoi maps. It is clear that the pictures with different degrees of generalisation are similar. Figure 17.4 proves this assumption. The bar plots of the information content (entropy) calculated using different methods are presented in Figure 17.5. We can see that the changes in generalised building map entropy are very close to pseudo-random map entropy changes. This means that the model of generalisation described maintains the geographical correctness and characteristics of the urban area.

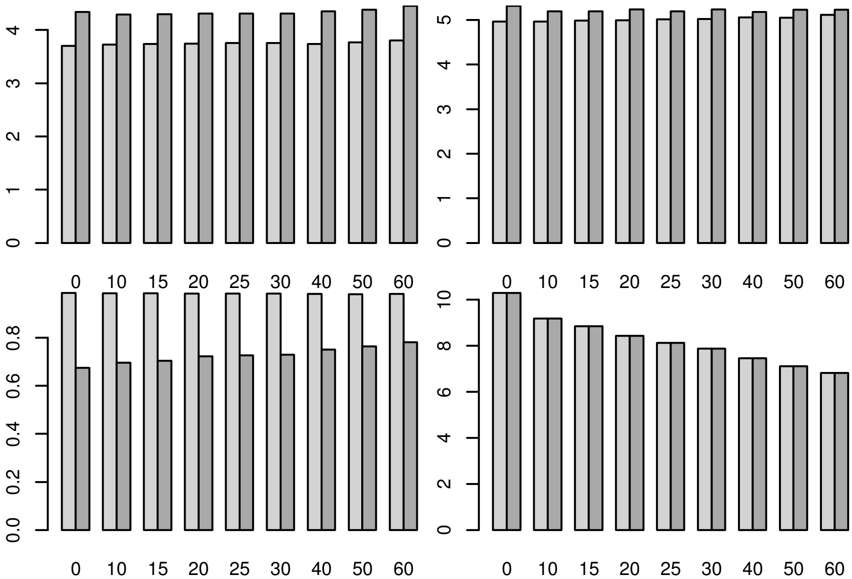


Fig. 17.5 Information content (entropy) of buildings layers (dark grey bars) and random Voronoi maps (light grey bars). The vertical axis is entropy, horizontal degree of generalisation (0-original buildings): the top left and top right) information content of the maps calculated using the CDD method using Eq. 4 for easting and northing coordinates, correspondingly; bottom left) normalised by number of polygons entropy of Voronoi regions (Eq. 6); bottom right) entropy of Voronoi neighbours

6 Compiling of 3D Scene

Noskov and Doytsher (2013b) presented an algorithm for compiling a 3D scene based on buffer zones. Buffer zones around the viewpoint were calculated, then quarters were selected from the hierarchy by zones, and spatial conflicts were resolved. This algorithm requires artificially defining buffer zone sizes and the correspondence between buffer zones and LODs. Our new approach is totally automatic and does not use distance from viewpoint to define the level of generalisation for some quarters. The approach is based on the information content of 3D scenes saved to raster file.

At this time, a database of the hierarchy of quarters and multiple building layers generalised in frames of the quarters was developed. The first (original) level of quarter hierarchy contains original buildings, generalised by pixel sizes of 10 and 15 meters; the second level corresponds to a 20 and 25 meters degree of generalisation; level 3 to 30 m, level 4 to 40m, level 5 to 50 and level 6 to 60m.

To reduce the number of quarters for processing, only those visible on 3D scene quarters were selected. We used a simple method which is demonstrated in Figure 17.6. The raster of a 3D scene was generated with relief, and then a raster with the relief overlaid by the spot of a quarter was calculated. If the two binary pictures were not same, a quarter was marked for further processing; otherwise, the quarter was excluded.

Each visible quarter on the 3D scene was processed to define the optimal level of generalisation. Quarters were ordered by distance from viewpoint; the nearest quarter was processed first. For each quarter, all levels of generalisation were considered. For each level of generalisation (including original buildings) a 3D scene's raster was calculated with resolution x,y containing only correspondent generalised buildings in the frame of the current quarter, and a 3D scene's raster with resolution x,y multiplied by k containing original buildings in the frame of the current quarter. We chose 650,340 for x,y and 5 for k; this was defined with respect to computer configuration. We then converted the raster scene to vector data; all pixels grouped with the same value formed vector polygons (Fig. 17.7). Analysis of 20 different quarters with respect to all degrees of generalisation was performed. Different ways of calculating entropy were tested. The measure which enabled us to define the optimal level of generalisation was found to be:

$$E = \left| \frac{Hg\_c - Horig\_c}{Hg\_0 - Horig\_0} \right|, \text{until} \quad E \le 7 \qquad (8);$$

All $H*$ parameters were calculated using Eq. 5; instead of Voronoi regions, areas of vectorised pixel groups were used. In the formula, *g* means the generalised buildings and normal 3D scene's raster resolution, *orig* means original buildings and high 3D scene's raster resolution, *c* means current region's frame, *0* means 0 degree of generalisation or original buildings. The expression $Hg\_0 - Horig\_0$ enables us to normalise entropy by eliminating the effect of pixel resolution differences. We use different resolutions of raster files because distant objects in a 3D scene can be depicted as 1 pixel objects (in which case it is impossible to calculate the real entropy of a sample raster); to calculate the entropy of a sample raster with original buildings we therefore needed to enlarge resolution. According to the equation presented above, a level with a maximum degree of generalisation with $E \le 7$ is appropriate for inclusion in the final generalised 3D scene. Value 7 was derived empirically. The *E* parameter could act as the coefficient of degree of the generalisation of a 3D scene.

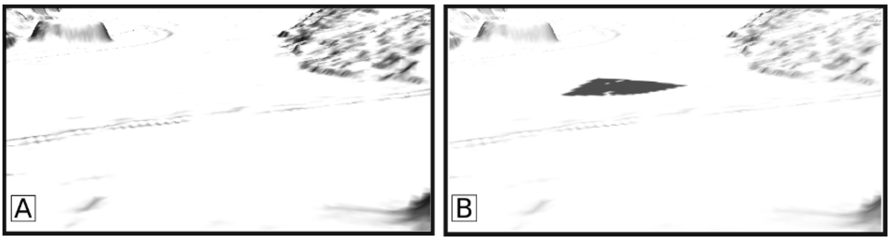


Fig. 17.6 Testing a quarter for visibility in a 3D scene: A) Empty scene with relief only; B) Scene with relief overlaid by the spot of a quarter

Fig. 17.7 Entropy of rasterised 3D scene: A) High resolution raster of original buildings; B) Normal resolution raster with generalised buildings; C) Vector polygon map of visible buildings' faces derived from B

The algorithm can be presented in pseudo code as follows:

```
hierarchy_of_ quarters
buildings            #Contains original buildings separated by quarters (horizontally) and by
resolution (vertically)
dem                  #Digital Elevation Model
quarters_of_1level=get_quarters(hierarchy_of_ quarters, level =1)
quarters_of_1level_of_scene=[]
for cur_quart  in quarters_of_1level{
```

```
       if draw_3D_as_PNG(dem, as.raster(cur_quart)) != draw_3D_as_PNG(dem){
           append_to_list(quarters_of_1level_of_scene, cur_quart)       }   }
function e(raster1,raster2){
      areas1=get_polygon_areas(rast_to_vector(rast1))
    areas2=get_polygon_areas(rast_to_vector(rast2))
    return abs(entropy(areas1)- entropy(areas(2)))   }
final_list=[]
for quart in quarters_of_1level_of_scene{
     zero_E=null
     selected_buildings_layer=null
     for resolution in [0,10,15,20,25,30,40,50,60]{
           hyerarchy_level=get_hlevel_by_resolution(resolution)
           cur_quart =  get_quarters(hierarchy_of_ quarters,  level=hierarchy_level, select_by_map=
quart)
           if cur_quart NotIntersectsWith final_list{
                  generalised_buildings=get_buildings(resolution=resolution, select_by_map=cur_quart)
                  original_buildings=get_buildings(resolution=0, select_by_map=cur_quart)
                  out_ras_resolution=650,340
                  rast_gen=draw_3D_as_PNG(generalised _buildings, resolution=out_ras_resolution)
                  rast_orig=draw_3D_as_PNG(original_buildings, resolution=out_ras_resolution*5)
                  E= e(rast_gen, rast_orig)
                  if resolution == 0 { zero_E =E}
                  if ( E<=7 ) { selected_buildings_layer= rast_gen }
                  else { break }            }     }
     appent(final_list, selected_buildings_layer)
  }
final_raster=draw_3D_as_PNG(dem, final_list)
original_buildings=get_buildings(resolution=0, select_by_map=cur_quart)
                  out_ras_resolution=650,340
                  rast_gen=draw_3D_as_PNG(generalised _buildings, resolution=out_ras_resolution)
                  rast_orig=draw_3D_as_PNG(original_buildings, resolution=out_ras_resolution*5)
                  E= e(rast_gen, rast_orig)
                  if resolution == 0 { zero_E =E}
                  if ( E<=7 ) { selected_buildings_layer= rast_gen }
                  else { break }            }     }
     appent(final_list, selected_buildings_layer)  }
final_raster=draw_3D_as_PNG(dem, final_list)
```

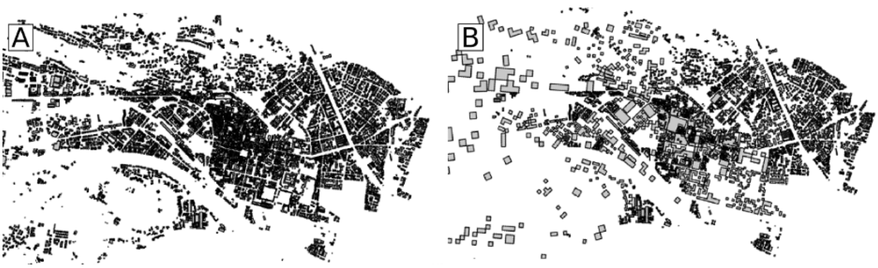


Fig. 17.8 2D footprints depicted on final 3D scene: original buildings (left); generalised buildings (right). View direction – from right to left side of the picture

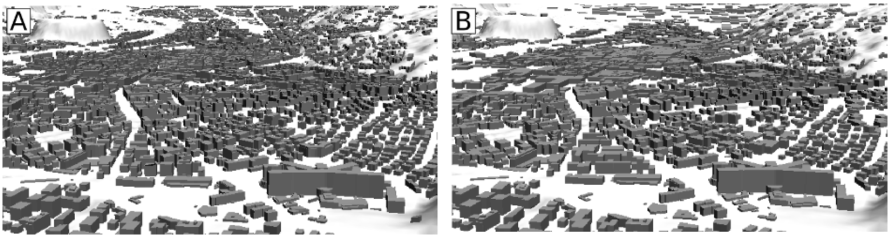


Fig. 17.9 Final 3D scenes: original scene (left); generalised scene (right)

7 Concluding Results

In Figure 17.8, the 2D footprints of the buildings depicted in the final 3D scenes are presented (see Figure 17.9). In contrast to approaches using distance to the object from a viewpoint as the main parameter for defining LODs, the information theory based method is more flexible; this method does not require defining the correspondence between distances from a viewpoint and LODs. According to Table 17.1, the number of nodes and the number of polygons decreased almost three times due to the generalisation, thus the speed of the rendering also decreased almost three times. In addition, the size of the final raster file in PNG format is decreased.

A raster-based cartographic generalisation approach was presented here. It is based on standard tools of rasterisation, vectorisation, region growing, classifying, and overlaying. The main advantages are: (1) the ability to simplistically and efficiently generalise buildings at different levels; (2) achieving variable and continuous level-of-detail of the buildings; (3) the generalised 3D model does not contain unreadable and overly detailed data of separate buildings; (4) the model maintains the geographical correctness and characteristics of the urban area; (5) the developed method helps to reduce computing time, and the computer resources required. The results of generalisation of buildings' footprints were evaluated. The proposed method of compiling 3D scene enables us to define the optimal LODs of objects automatically, and takes into account the information content of a 3D scene.

The method and the process were developed by using a standard PC (DELL Vostro 3550), 4 processors: Intel® Core™ i3-2310M CPU @ 2.10GHz, with 1.8 GB Memory. In addition, Debian GNU/Linux 7 operating system, GRASS GIS 7, Bash and R programming languages were used. Calculation of raw quarters, a quarter's hierarchy and multiple LODs for the city area takes several hours. A parallel computing approach for Graphic Cards (OpenCL application) could significantly reduce the required time. Defining optimal LODs for one viewpoint requires approximately 10-15% of the time required for the raw quarter calculations. The time could again be reduced by simultaneously calculating 3D scene parameters for several viewpoints and using them in artificial neural networks for fast optimally defined LODs.

Table 17.1 Comparison of numerical parameters describing the advantages of implemented generalisation

	Number of nodes	Number of polygons	Speed of 3D scene generating, seconds	Size of generated PNG raster file, kilobytes
Original building layer	114,197	12,527	10.2	120
Generalised building layer used for 3D	38,343	4,552	3.4	99

References

Battersby SE, Keith C (2003) Quantifying Information in Line Generalisation. Department of Geography, University of California at Santa Barbara, CA 93106 USA, Proceedings of the 21st International Cartographic Conference (ICC), Durban, South Africa

Bjorke JT (1996) Framework for entropy-based map evaluation. Cartography and Geographical Information Systems 23: 78-95

Clarke, Keith C, Battersby SE (2001) The Coordinate Digit Density function and Map Information Content Analysis. Proceedings of the American Congress on Surveying and Mapping Annual Meeting, Las Vegas, NV

Forberg A (2007) Generalisation of 3D Building Data Based on a Scale-Space Approach. ISPRS Journal of Photogrammetry & Remote Sensing 62:104-111

Glander T, Döllner J (2009) Abstract representations for interactive visualization of virtual 3D city models. Computers, Environment and Urban Systems 33: 375–387

Guercke R, Götzelmann T, Brenner C, Sester M (2011) Aggregation of LoD 1 building models as an optimization problem. ISPRS Journal of Photogrammetry and Remote Sensing 66: 209–222

He S, Moreau G, Martin J (2012) Footprint-Based 3D Generalisation of Building Groups for Virtual City. GEOProcessing 2012: The Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services, Valencia, Spain

Kada M (2002) Automatic Generalisation of 3D Building Models. Proceedings of the Joint International Symposium on Geospatial Theory, Processing and Applications, Ottawa, Canada

Knopfli R (1983) Communication theory and generalisation. In: Graphic Communication and Design in Contemporary Cartography, ed. by D.R.F. Taylor, 177-218, New York, John Wilyey & Sons Ltd

Li Z, Yan H, Ai T, Chen J (2004) Automated building generalisation based on urban morphology and Gestalt theory. International Journal of Geographical Information Science 18 (5): 513-534

Li Z, Huang P, (2002) Quantitative measures for spatial information of maps. International Journal of Geographic Information Science 16(7): 699-709

Neumann J (1994) The topological information content of a map / An attempt at rehabilitation of information theory incartography. Cartographica, 31(1): 26-34.

Noskov A, Doytsher Y (2013a) Urban perspectives: A raster-based approach to 3D generalization of groups of buildings. GEOProcessing 2013: The Fifth International Conference on Advanced Geographic Information Systems, Applications, and Services, Nice, France

Noskov A, Doytsher Y (2013b) Hierarchical quarters model approach toward 3D raster based generalization of urban environments. International Journal on Advances in Software 6 (3&4): 343-353

Shannon CE (1948) A mathematical theory of communication. The Bell System Technical Journal, 27: 379-423 & 623-656

Sukhov VI (1970) Application of information theory in generalization of map contents. International Yearbook of Cartography, 10: 41-47

Thiemann F (2002) Generalization of 3D Building Data. Proceedings of Symposium of Geospatial Theory, Processing and Applications, Ottawa, Canada

Trapp M, Glander T, Buchholz H (2008) 3D Generalization Lenses for Interactive Focus + Context Visualization of Virtual City Models. Proceedings of the 12th International Conference Information Visualization, London, UK

Xie J, Zhang L, Li J (2012) Automatic simplification and visualization of 3D urban building models. International Journal of Applied Earth Observation and Geoinformation 18: 222-231

Quantification and Mapping of Flood Regulating Ecosystem Services in Different Watersheds – Case Studies in Bulgaria and Arizona, USA

Kremena Boyanova*, Stoyan Nedkov, Benjamin Burkhard

Abstract There is great need for accurate and practical methods to assess the conditions of ecosystems, and the possible results of their interaction with social systems. The generation and interpretation of quantitative data for ecosystem service analysis is still not well established. Ecosystem service analyses demand an interdisciplinary approach that integrates knowledge with a high variety, and manifold verifications, of models and data. Maps seem **to** be the most preferable tool for the visualisation of results, being a comprehensive and intuitive tool for communication between decision makers and the general public. The following chapter presents an application and the verification of an approach for the quantification of flood regulating ecosystem services by using results from the watershed hydrological model KINEROS and the AGWA tool (Nedkov and Burkhard 2012). It is applied in six watersheds - three in Bulgaria and three in Arizona, USA, in order to check its reliability in case studies with differing geographic characteristics. The model results are used to define the capacities of the land cover classes in the different watersheds and to prepare flood regulating supply capacity maps. Capacities for flood regulation differ within the case studies and their land cover classes. Forests still show generally high capacities in both Bulgaria and Arizona, while grasslands and pastures in Bulgaria show higher capacities for flood regulation than in Arizona. The maps can provide valuable information for sustainable environmental management.

K. Boyanova*, S. Nedkov
Department of Geography, National Institute of Geophysics, Geodesy and Geography, Bulgarian Academy of Sciences
Akad. G. Bonchev str. bl.3, 1113 Sofia, Bulgaria
*corresponding author
e-mail: kbboyanova@gmail.com, Phone: (+3592) 979-3360, Fax: (+3592) 979-3360

B. Burkhard
Institute for Natural Resource Conservation, University of Kiel
Olshausenstr. 40, 24098 Kiel, Germany

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_18,

1 Introduction

1.1 What are Ecosystem Services?

Until now, science and practice have mostly been looking at environmental and social systems separately, as independent and self-organised units with open flows of energy, water and matter. The Ecosystem Services (ES) concept provides a common ground for the analysis of the social system and the ecosystem as parts of a bigger system, the so called human-environmental system, where the relationships between humanity and natural resources are much better established. Such a concept provides a better understanding of possible human contributions to the health of the whole system. In the years since the establishment of the ES concept, various definitions have been presented, but they significantly overlap and supplement each other. According to the "Salzau Message" on Sustaining Ecosystem Services and Natural Capital, *ecosystem services* are "the contributions of ecosystem structure and function – in combination with other inputs – to human well-being" (Burkhard et al. 2012a). This definition incorporates both the natural and social systems and their interactions, since the "combination with other inputs" refers to the anthropogenic inputs in the natural systems, which nowadays have increased considerably. Ecosystem services are classified in three main groups – regulating, provisioning and cultural ES. Regulating services contribute to human well-being through the regulation of natural processes, such as climate regulation, or nutrient regulation. Provisioning services are materials and products that humans obtain from nature and use, for instance, for nutrition or as energy. Cultural services are the non-material, intangible benefits that humans obtain from ecosystems, such as recreation, spiritual and religious experience, inspiration, or education. Flood regulation is a regulating service that incorporates parts of two other regulating services – water flow regulation and natural hazard protection (Kandziora et al. 2013a).

1.2 Quantifying Ecosystem Services

There are many methods of "measuring" ecosystem services, and they depend on service and area specifics, research questions and objectives, or data availability. For better understanding of ecosystems and the ways they contribute to human well-being through services, it is necessary to be familiar with the quality and quantity of those services.

The development of methods to quantify ecosystem services has been recognised as an important step for the successful application of the concept in management and decision making (de Groot et al. 2010). The bio-physical elements and processes, as well as their interrelations, determine the potential supply and flow of various ecosystem services. Cultural services are more specific, since their benefits are more qualitative than quantitative and most of their measures are highly subjective. Very often cultural services have only a local or regional supply, which is, for example, the case with religious and spiritual services. Futhermore, supply potential should be differentiated from actual supply, for certain services. This distinction would indicate the difference between the total quantity of a service that is present in an area (potential supply) and the quantity of that service actually used by the human society (actual supply). It is easier to differentiate those values for provisioning services than for regulating services. It is, for example, extremely hard to separate the portions of climate regulation services that are actually used from the total supply of these specific services. The same is the case with flood regulation – the potential and the actual supply overlap and represent the total capacity for the service provision.

1.3 Hydrological Modelling for the Quantification of Ecosystem Services

Hydrological modelling is a very broad field in the natural sciences, including knowledge about hydrological processes in ecosystems, the elements that are involved in them and their interactions. Water is integrated in many ways in our lives, being the most essential element of existence. The task of analysing and better understanding all components of the water cycle is extremely challenging. Still, scientists try to do so by simplifying models that reproduce processes in nature in their efforts to understand and use them in a better and more sustainable way. According to a review by Vigerstol and Aukema (2011), both the application of hydrological modelling and ecosystem service models find broad application for water-related ecosystem service assessment. Depending on research needs and data availability, the application of hydrological models is preferable in cases where sufficient data is available and related expertise is needed. Respective tools incorporate the knowledge and professional experience of water resource specialists.

1.4 Mapping Ecosystem Services and its Role in European Commission Legislation

The concept of ecosystem services, together with ES mapping, is highly applicable to environmental management, policy and decision making. Its incorporation in European Commission legislation has increased rapidly (EC Biodiversity Strategy 2020[2], General Union Environment Action Programme to 2020[3]). In Action 5 Target 2 of the Biodiversity Strategy 2020, the European Commission demands that member states "*improve knowledge of ecosystems and their services in the EU (Action 5) – the member states shall map and assess the state of ecosystems and their services in their national territory by 2014, assess the economic value of such services, and promote the integration of this values into accounting and reporting systems at EU and national level by 2020*".

Mapping ES has broad application potential since it is an extremely valuable method for visual representation of qualitative and quantitative spatial data. Nevertheless, the various generalisation techniques are necessary steps in the mapping process because it is (still) impossible to represent the complexity of nature and anthropogenic impacts and changes in a single map.

There are many variables in the process of mapping ecosystem service supply, such as the mapped service, scale, indicators, or methods and models used (for review see: Crossman et al. (2013), Martinez-Harms and Balvanera (2012). Quantifying and mapping the supply capacities of different services is necessary for analysing the trade-offs between different ES. However, it can be difficult to make comparisons between elements measured or modelled in different quantification units, or at different scales or resolutions. In this chapter, the "matrix method" of mapping the capacity of different land cover types to supply ecosystem services (Burkhard et al. 2009) was applied. It is a land cover based assessment approach in which land cover classes are used as basic spatial units. Land cover classifications and the respective maps are usually familiar to decision-makers and easy to understand by the general public.

For the assignment of capacity classes, various data, such as quantitative data (e.g. from modelling), interview results, expert-based assessments and others, can be used. In the case study presented, quantitative data from hydrological modelling was applied. The results were used for mapping the flood-regulating supply capacities of the land cover classes in the different watersheds. Capacities for flood regulation differ within the case studies and their land cover classes. Forests, for example, show generally high capacities in both Bulgaria and Arizona, while grasslands and pastures in Bulgaria show higher capacity for flood regulation than in Arizona.

[2]http://ec.europa.eu/environment/nature/biodiversity/comm2006/2020.htm
[3]http://ec.europa.eu/environment/newprg/proposal.htm

2 Material and Methods

2.1 Case Study Areas

2.1.1 Bulgaria

The approach was applied to three watersheds in northern Bulgaria: Ravna, Vidima and Yantra (Fig. 18.1) in order to assess the capacities of different land cover types to supply flood regulation ecosystem services. The watersheds are located along the northern slopes of the Balkan Mountains. The Ravna watershed is the smallest with an area of 2578.6 ha. It is located upstream of the village of Boikovets and has an elevation ranging from 680 to 1479m. The Vidima watershed is east of Ravna, upstream of the town of Apriltsi and covers an area of 7677.9ha with an elevation ranging from 505 to 2375m. The biggest watershed is Yantra, upstream of the town of VelikoTarnovo, covering an area of 28627ha, which makes it the biggest of the three areas. It has an elevation ranging from 357 to 1440m. All three watersheds have a temperate-continental climate which is characterised by relatively cold winters and warm summers. The annual precipitation varies from 750-800mm in the lower parts to 1100-1200mm in the higher mountains. Extreme precipitation events are typical for these areas. The flood hazard in the area is very high (Nedkov and Nikolova 2006, Nikolova et al. 2007, Nikolova et al. 2008). The flood regulating capacities of the ecosystems are therefore of significant importance to human well-being in the area.

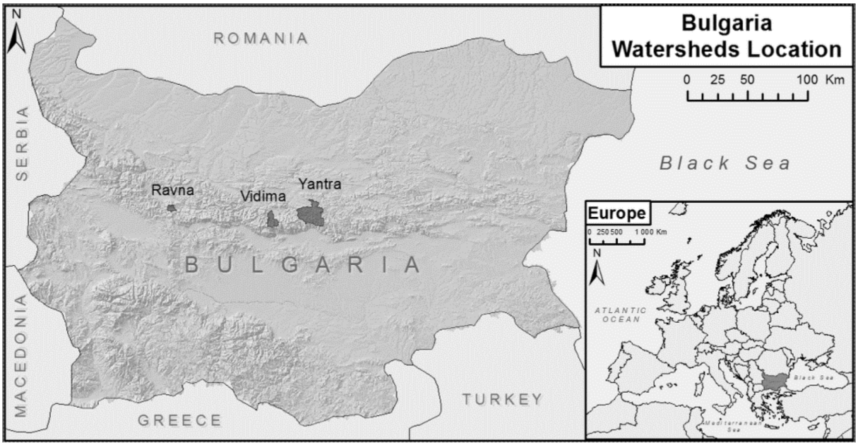


Fig. 18.1 Location of the three case study watersheds in Bulgaria

2.1.2 Arizona

The San Pedro watershed is located in both Mexico and the USA. The river originates in Sonora, Mexico and flows northwards into south eastern Arizona, USA (Fig. 18.2). The watershed has a semi-arid climate characterised by very hot summers and mild to warm winters. The annual precipitation ranges from 314 to 635mm. Rain often occurs in the form of thunderstorms causing floods in July and August with considerable local damage[4]. Three sub-watersheds were chosen in the US part of the San Pedro watershed. They are called *A, B* and *C* here, since their names were not provided in the data source. Sub-watershed *A* is the biggest of the three sub-watersheds with an area of 9213.3ha and an elevation ranging from 1255 to 2605m. Sub-watershed *B* has a common border with sub-watershed *A*. It covers an area of 3207.4ha and has an elevation ranging from 1250 to 2458m. The upper parts of both sub-watersheds are located in the Hauchuca Mountains. The smallest sub-watershed *C*, located in Mule Mountains, covers an area of 2989.3ha and its elevation ranges from 1426 to 2208m.

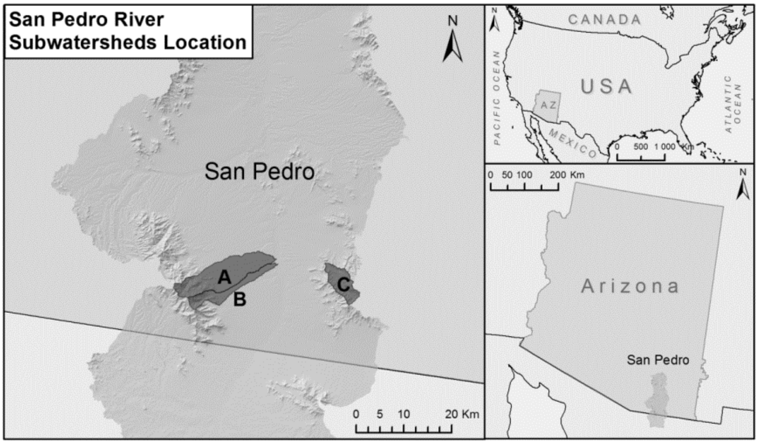


Fig. 18.2 Location of the three case study sub-watersheds in Arizona, USA

2.2 Hydrological Model KINEROS and AGWA Tool

The GIS-based AGWA (Automated Geospatial Watershed Assessment; Miller et al. 2002) tool and its constituent model KINEROS (KINematic Runoff and EROSion model, Semmens et al. 2005) were applied for the modelling of hydrological processes in the chosen watersheds. KINEROS is an event-oriented, physically-based model, describing the processes of interception, infiltration, surface runoff and erosion from small agricultural and urban watersheds. It can be used to deter-

[4]www.wrcc.dri.edu/narratives/ARIZONA.htm

mine the effects of various artificial features such as urban developments, small water storage reservoirs, or lined channels of flood hydrography and sediment yield[5].

2.2.1 Input Data and Calibration

Bulgaria

The KINEROS model was run for the test watersheds in Bulgaria using a 30m Digital Elevation Model (DEM) derived from 1:25 000 topographic maps by manual digitalisation, digital soil map data (FAO/UNESCO 2003) and CORINE Land Cover data (CLC) (Bossard et al. 2000). The geometric accuracy of the CORINE data is about 100m which is too coarse for the model requirements, so additional interpretations of Landsat ETM+ satellite images with 30m resolution was taken in July 2006 and orthophoto maps with 60 cm resolution and 5 m geometric accuracy were performed for the watersheds. The interpretation was conducted manually using the standard CORINE land cover nomenclature. The resulting land cover data is more detailed. For example, the Ravna watershed contains 11 land cover classes and 65 separate spatial units (individual vector polygons) in the original CLC2000 map, but has 14 land cover classes and 449 units after the additional interpretation. Appling input data with higher accuracy increases the precision of the hydrological modelling results, providing more representative data for the ecosystem service capacity assessment. More information on the relationships between input data accuracy and ecosystem service capacity assessment can be found in Kandziora et al. (2013b). The calibration for the Ravna watershed was performed for an event that occurred on May 26, 2005 with a precipitation of 59mm and a peak flow of 7.6m^3/sec. The data for Yantra was calibrated against an event on July 6, 1991 with a precipitation of 48mm and peak flow of 43m^3/sec. The adjustments for the Yantra event were used for the calibration of an event in the Vidima watershed, because data for this specific area was lacking.

Arizona

For the purpose of this research, freely available data from the tutorials of the United States Department of Agriculture (USDA) hydrologic model KINEROS (KINematic Runoff and EROSion) model was used[6]. The San Pedro subwatersheds were delineated using a 30m DEM derived from USGS Digital Eleva-

[5]http://www.tucson.ars.ag.gov/kineros/

[6]http://www.tucson.ars.ag.gov/agwa/index.php?option=com_jdownloads&Itemid=&view=viewdownload&catid=20&cid=223

tion Model data[7]. They were parameterised using STATSGO soil data derived from the USDA soil dataset and the North American Landscape Characterisation (NALC) dataset with 60m resolution derived from Landsat TM data (Scott 2005). The model was run for a 100 year one hour storm event from the AGWA database.

2.3 Capacity Assessment of Flood Regulating Ecosystem Services

For a proper assessment of the capacity of different land cover types to supply flood regulating ecosystem services, the choice of suitable indicators plays a significant role. One of the most basic scientific demands on good indicators is that they represent a clear proof of relevant cause-effect relations (Kandziora et al. 2013a). For the assessment of the flood regulating capacities in the test watersheds, infiltration, peak flow and surface runoff were chosen from the hydrologic parameters provided by the model. The 'relevant cause-effect relationship' between these parameters and flood regulation lies in the dependence of a flood event's magnitude on peak flow, infiltration and surface runoff. For example, the combination of high infiltration, low peak flow and low surface runoff in a certain area decreases the possibility of hazardous flood events due to the ability of the area to "capture" water. Based on this assumption, the following relationships between these three parameters and their flood regulating capacity have been developed (Table 18.1):

Table 18.1 Relation between hydrologic parameters and flood regulation capacity

Parameter	Value	Flood regulation capacity
Infiltration	high	high
Peak flow	low	high
Surface runoff	low	high

Assessment for the Bulgarian case studies was made using the CORINE land cover classification system and the detailed CLC2000 dataset, which was also used as input data for the hydrologic model. The NALC classification system and the NALC73 dataset were used for assessing the sub-watersheds in Arizona. Both CORINE and NALC data is freely available. Even though the two classification systems differ, a comparison based on the basic characteristics of the land cover classes was possible.

[7]http://edcwww.cr.usgs.gov/doc/edchome/ndcdb/ndcdb.html

2.4 Mapping the Capacity for Flood Regulation of the Different Land Cover Types

As mentioned earlier, the ES "matrix" method developed by Burkhard et al. (2009) was applied for the mapping of flood regulation capacities of the different land cover types in the study areas. The method uses a relative scale from 0 to 5 with 6 different ES supply classes. The developed colour legend simplifies the mapping process and provides an illustrative visualisation of the modelling results. The colour chosen to illustrate the areas with lowest capacity (0 – rosy colour - no relevant capacity) contrasts with the colours of all the other capacities, which are in the green palette (1 - grey green - low relevant capacity, 2 - light green - relevant capacity, 3 - yellow green - medium relevant capacity, 4 - blue green - high relevant capacity and 5 - dark green - very high relevant capacity (Burkhard et al. 2009)).

Burkhard et al. (2012b) showed how maps of supply capacities and the colour scale legend from 0 to 5 can further be combined with maps of demand capacities with a colour scale legend from 0 to -5, resulting in ecosystem service budget maps.

The scoring is based on the data available for the study area. In this sense, 0 indicates the lowest supply of a certain service in the area and 5 is the highest supply in the same area. The values in the six classes are separated based on equal intervals. This guarantees that, for example, class 4 supplies twice as much capacity as class 2 and makes the classes more comprehensive. This clarification is important for the flood regulating ecosystem service capacity assessments in Bulgaria and Arizona, which were made based on quantitative model data. A multistage GIS analysis approach, after Nedkov and Burkhard (2012), was applied for the final compilation of the flood regulation ecosystem service supply maps:

1. The model values for each hydrologic parameter in the respective watersheds were divided into six equal intervals[8] (one for each 0 to 5 class) (Table 18.2) – this determines the spatial distribution of the capacity classes, based on the relations from Table 18.1 (for example areas with low infiltration values of 6.59 - 7.93 mm in the Ravna watershed have a flood regulation capacity of 0).
2. These results were overlaid with the land cover data.
3. The spatial extension of each land cover class within the different capacity classes was calculated in order to determine how much of a certain land cover class falls in each capacity class (measured for each hydrologic parameter separately).

---

[8]The results from the hydrological modelling are divided into classes using the equal interval classification method, while in Nedkov and Bukhard (2012) the quantile classification method was used. This change in the approach explains differences in the results for the same areas in the two studies. Further analyses of those differences are needed.

4. The average values of these areas were calculated based on the results for each parameter (for example, if land cover class “Pastures” was measured as having an area of 195.9ha in infiltration capacity class 5, an area of 195.9 in peak flow capacity class 5 and an area of 193ha in surface runoff capacity class 5, then the area of pastures in the total flood regulation capacity 5 was 195.3ha).
5. The capacity class containing the largest area of a certain land cover class was detected.
6. The detected capacity class (from the previous example – if the area of “Pastures” in class 5 was the largest area for all the 6 classes, then pastures have flood regulating capacity 5) was ascribed to the land cover class.

These results were used for the compilation of flood regulating capacity maps of the different land cover types, applying the same colour legend as shown in Table 18.2.

The land cover maps of the case study areas in Bulgaria were made using the CORINE land cover three level classification system and the RGB colour legend provided by the dataset. The land cover maps of the case studies in Arizona are made using the NALC one level classification system and the RGB colour legend that is provided with the data.

Table 18.2 Ranges of the model result values for the indicators of flood regulation ecosystem service supply

Capacity class	Bulgaria			San Pedro, Arizona, USA		
	Ravna (2578.6 ha)	Vidima (7677.9 ha)	Yantra (28627.0 ha)	A (9186.8 ha)	B (3190.1 ha)	C (2980.0ha)
Infiltration	Model results (mm)					
0	6.59 - 7.93	34.15 - 36.24	25.78 - 29.15	17.40 - 19.27	17.60 - 19.49	13.71 - 14.13
1	7.94 - 9.28	36.25 - 38.33	29.16 - 32.51	19.28 - 21.13	19.50 - 21.38	14.14 - 14.55
2	9.29 - 10.63	38.34 - 40.42	32.52 - 35.88	21.14 - 23.00	21.39 - 23.27	14.56 - 14.97
3	10.64 - 11.97	40.43 - 42.51	35.89 - 39.24	23.01 - 24.86	23.28 - 25.15	14.98 - 15.39
4	11.98 - 13.32	42.52 - 44.60	39.25 - 42.60	24.87 - 26.73	25.16 - 27.04	15.40 - 15.81
5	13.33 - 14.67	44.61 - 46.69	42.61 - 45.97	26.74 - 28.59	27.05 - 28.93	15.82 - 16.23
Peak flow	Model results (mm)					
0	25.17 - 21.30	17.41 - 14.81	22.95 - 19.52	45.21 - 38.55	66.34 - 57.68	73.49 - 71.61
1	21.29 - 17.41	14.80 - 12.20	19.51 - 16.07	38.54 - 31.87	57.67 - 49.00	71.60 - 69.72
2	17.40 - 13.52	12.19`- 9.59	16.06 - 12.63	31.86 - 25.20	48.99 - 40.32	69.71 - 67.82
3	13.51 - 9.64	9.58 - 6.98	12.62 - 9.18	25.19 - 18.53	40.31 - 31.64	67.81 - 65.93
4	9.63 - 5.75	6.97 - 4.37	9.17 - 5.74	18.52 - 11.85	31.63 - 22.96	65.92 - 64.04
5	5.74 - 1.85	4.36 - 1.75	5.73 - 2.28	11.84 - 5.17	22.95 - 14.27	64.03 - 62.14
Surface runoff	Model results (mm)					

0	8.48 - 7.10	13.74 - 11.55	22.10 - 18.56	12.82 - 10.98	20.97 - 19.08	25.72 - 25.22
1	7.09 - 5.70	11.54 - 9.35	18.55 - 15.01	10.97 - 9.13	19.07 - 17.18	25.21 - 24.70
2	5.69 - 4.31	9.34 - 7.16	15.00 - 11.46	9.12 - 7.28	17.17 - 15.29	24.69 - 24.18
3	4.30 - 2.92	7.15 - 4.96	11.45 - 7.91	7.27 - 5.43	15.28 - 13.39	24.17 - 23.66
4	2.91 - 1.52	4.95 - 2.76	7.90 - 4.36	5.42 - 3.58	13.38 - 11.49	23.65 - 23.14
5	1.51 - 0.12	2.75 - 0.55	4.35 - 0.80	3.57 - 1.71	11.48 - 9.58	23.13 - 22.62

3 Results

3.1 Land Cover Type Capacities to Supply Flood Regulating Ecosystem Services

Repeating the approach presented above for each of the six watersheds provides an opportunity for comparative results analyses (Table 18.3). In all watersheds studied, in Bulgaria as well as in Arizona, the results for land cover classes with a spatial extension of less than 15ha were considered unrepresentative. In such small areas, model results are strongly influenced by surrounding land cover types and cannot be modelled correctly. Where possible, their capacity values were taken from results for the other watersheds in the same climate zone (semi-arid or temperate-continental respectively), where they were present with areas larger than 15ha. If this information was also not available, their capacities were based on expert opinion.

Table 18.3 Flood regulating ecosystem service supply capacities of the different land cover classes (empty fields indicate that the land cover class was not present in the respective watershed)

Land cover class	Capacities by watershed		
CORINE	Ravna	Vidima	Yantra
112 Discontinuous urban fabric	2	3	3
121 Industrial or commercial units	-	3	3
122 Road and rail networks	5*	-	5
141 Green urban areas	-	-	3
142 Sport and leisure facilities	-	-	5
211 Non-irrigated arable lands	2	3*	3
222 Fruit trees and berries	5*	5	5
231 Pastures	2	5	5
242 Complex cultivation patterns	-	5	5

243 Agriculture and natural vegetation	2	5	5
311 Broad-leaved forests	3	5	5
312 Coniferous forests	5	5	5
313 Mixed forests	3	5	5
321 Natural grasslands	5	3	5
322 Moors and heathland	5	5	5
324 Transitional woodland-shrub	3	4	5
332 Bare rocks	3	3	3*
333 Sparsely vegetated areas	3*	3	5
NALC	A	B	C
1 Forest	5	5	-
2 Oak woodland	5	4	2
3 Mesquite woodland	2	3	-
4 Grassland	5	3	3
5 Desert shrub	4	3	3
6 Riparian	-	-	2**
8 Urban	1	1*	-

* Capacity values were taken from the results in the other watersheds in the same climate zone (semi-arid or temperate-continental); see text for further explanation.
** Capacity values are based on expert assessment.

3.1.1 Comparative Results from the Case Studies in Bulgaria

Looking at the results for the three watersheds in Bulgaria it is clear that the Ravna watershed shows most significant difference to the other two watersheds. Ravna is also the watershed with the smallest area, even if the difference between the areas of Ravna and Vidima is smaller than the difference between Vidima and Yantra. Vidima and Yantra indicate similar capacities for flood regulation in twelve land cover classes. Only two land cover classes are present in all three watersheds and show the same capacities: *Coniferous forest* and *Moors and heathland.* Ravna and Vidima have the same results for three land cover classes: *Coniferous forest, Moors and heathland* and *Bare rocks*, and Ravna and Yantra also have the same results for three land cover classes: *Coniferous forest, Natural grasslands,* and *Moors and heathland.* In general, the three watersheds show relatively high capacities for flood regulation, and capacity classes 0 and 1 are not present in any of them. The number of land cover classes with very high capacity for flood regulation is high. Most significant are the areas with natural vegetation and agriculture, as also shown by Stürck et al. (2014). The urban areas show mainly moderate capacities and *Roads and rail networks* show very high capacities, probably because of their engineering characteristics.

3.1.2 Comparative Results from the Case Studies in Arizona

The overview of the results for the three sub-watersheds in Arizona indicates that the smallest – *C,* shows the lowest capacities for flood regulation and represents only four land cover types. There is no land cover class that is present in all three sub-watersheds and has the same capacity class. Sub-watersheds *A* and *B* show the same capacities for one land cover class: *Forest,* and *B* and *C* show same capacities for two land cover classes: *Grassland* and *Desert shrub*. Capacity class 0 is not present in any of the sub-watersheds. The land cover class *Riparian* has no representative areas in any of the sub-watersheds. The highest capacities are indicated for *Forest* and *Oak woodland.* From the classes that represent natural vegetation, *Mesquite woodland* shows the lowest capacities. The urban areas show low and moderate capacities for flood regulation.

3.1.3 Comparative Results from the Case Studies in Bulgaria and Arizona

In both Bulgaria and Arizona, forests show high capacities for flood regulation while urban areas show relatively low capacities in the respective areas – low in Arizona and medium in Bulgaria. Grasslands and pastures in Bulgaria show higher capacities for flood regulation than in Arizona. It is notable that in both Bulgaria and Arizona the watersheds with the smallest area (under 3000 ha) show the greatest differences in flood regulation capacity compared to the other watersheds in the respective country and the generally lower capacities for flood regulation indicate its incapacity for self-regulation of extreme flood events.

3.2 Maps of Flood Regulating Ecosystem Service Supply

3.2.1 Temperate-Continental Climate Case Studies in Bulgaria

The maps of the flood regulating capacity of the different land cover classes were made according to the colour legend provided by Burkhard et al. (2009) and also applied by Nedkov and Burkhard (2012). They are related to the colours from Table 18.3. The resulting maps (Fig. 18.3) are arranged in a manner that provides easy comparison between the two thematic maps in the respective areas.

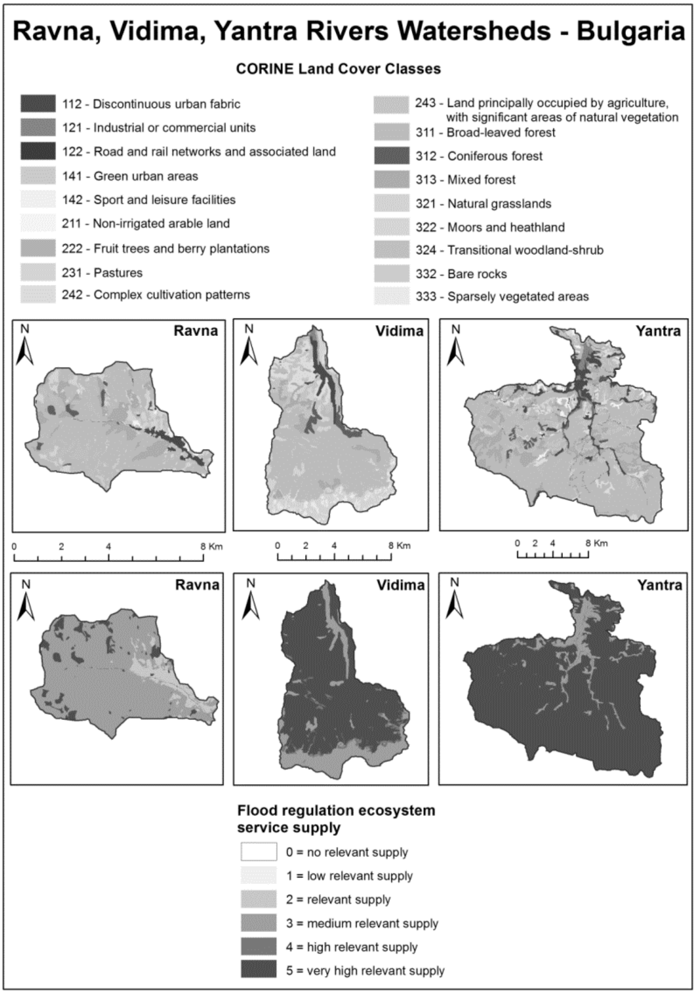


Fig. 18.3 Land cover maps (top) and maps of flood regulating ecosystem service supply capacities (bottom) in the three study areas in Bulgaria

3.2.2 Semi-Arid Climate Case Studies in Arizona

The maps of the flood regulating capacity of the different land cover classes in the Arizona case study were made using the same method as for the Bulgarian maps. This enables comparative analyses of the results, independent from the available land cover datasets in the area, their climate or geographic location. The resulting maps are again arranged in a manner enabling easy comparison between the land cover and the flood regulation capacities shown in the thematic maps in the respective areas (Fig. 18.4).

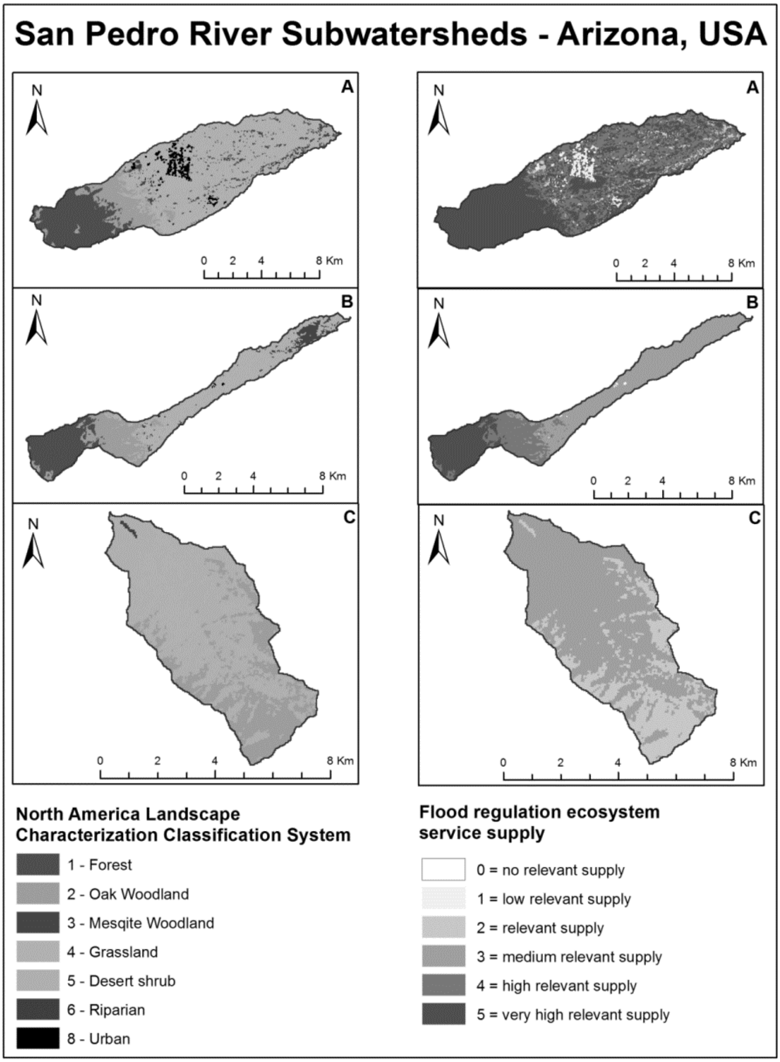


Fig. 18.4 Land cover maps (left) and maps of flood regulating ecosystem service supply capacities (right) in the three study areas in Arizona, USA

4 Discussion

Although the areas in Arizona and Bulgaria have significant differences in their climate, location and many other factors, the results of the flood regulation assessment show significant similarities. Through the application of the assessment approach in such different areas, some basic assumptions can be made:

- Watersheds with areas smaller than 3000ha do not seem to be representative enough for the application of the approach.
- Forests, in general, show high and very high capacities for flood regulation, both in the semi-arid and the temperate-continental climate zone. Mixed forests show especially high capacities.
- The capacity of shrublands and grasslands should not be underestimated, especially in semi-arid areas where the water uptake of forests and/or agricultural areas may not be acceptable because of regularly occurring droughts and generally insufficient water resources.

The high capacities of forests for water retention have been recognised in multiple studies (Biao et al. 2010, Bradshaw et al. 2007, Laurance 2007). Flood regulating capacities for the whole diversity of land cover classes have been analysed in various studies at different scales (Guo and Gan 2002, Nedkov and Burkhard 2012, Stürck et al. 2014)), corresponding to the results derived from this study. Nevertheless, the differentiation between the flood regulating capacities of the different natural vegetation types and agricultural areas was not clearly distinguishable. This may be influenced by outlier values in the model results, which can shift the assessment. The maps of flood regulating ecosystem services provide a clear picture of areas with high and low supply capacities and facilitate the choice of preferable land use practices for flood risk mitigation in areas troubled by flood events.

Scoring land cover capacities to supply ecosystem services enables further trade-off analyses. Ecosystem services are normally measured in different biophysical units, which complicates the comparison between them. The application of an integrative capacity assessment makes comparisons possible. The maps of the capacities for ecosystem service provision show the spatial distribution of areas with high and low capacity. This also provides a spatial overview of the possible directions of ecosystem service flows within an area, which is usually from areas with high capacities to areas with low capacities. Maps of different services in the same area may show spatial correlations between these services.

5 Conclusions

Even if the results show relatively high capacities for flood regulation, significant flood events still occur in all areas. The supply of flood regulation is generally lower than the demand and further measures for the prevention of flood hazards should be applied. A possible future step in the development of the assessment approach will be the elaboration of future scenarios of land cover changes and management solutions in order to assess their influence on the regulation of hazardous flood events. An awareness of possible solutions to such problems, using the natural regulating function of the environment, will increase the chances of a sustainable future.

There are many uncertainties in the approach presented, related, among other things, to input data quality and quantity, model shortcomings, indicator selection and map generalisation. Further analyses integrating information on soil types, better elevation data, socio-economic information and further biophysical parameters can be conducted for more detailed assessments of the different human-environmental systems. Nevertheless, even if using “only” land cover classes as basic spatial units for assigning flood regulating capacities, the approach provides clear patterns of the distribution of areas with high and low capacities in the resulting maps. In addition to this information, areas of interest for further and more detailed analyses can also be selected. The identification of areas with high/low capacities for flood regulation ecosystem service supply delivers information that is highly relevant for landscape planning and environmental management. An appropriate application of the respective information can help to develop more site-specific land use strategies towards more sustainable natural resource management. The approach presented and the conducted results may, for example, be applied in the implementation of the European Union (EU) Floods Directive[9] that was accepted in 2007 by all EU member countries. The aim of the directive is to provide guidelines for the reduction and management of the risks that floods pose to human health, the environment, cultural heritage and economic activity. Flood regulation assessment is one example where ecosystem service mapping and cartography can be of significant use to the society.

Acknowledgments The work for this study has been supported by the EU 7th Framework Program project SWAN (Sustainable Water ActioN; Grant agreement no: 294947) and by the "National, European, and Civilizational Dimensions of the Culture – Language – Media Dialogue" Programme of the "Alma Mater" University Complex for the Humanities at Sofia University "Saint Kliment Ohridski ", funded by the Bulgarian Ministry of Education, Youth, and Science Scientific Research Fund.

Reference

Biao Z, Wenhua L, Gaodi X, Yu X (2010) Water conservation of forest ecosystem in Beijing and its value. Ecological Economics 69: 1416-1426

Bossard M, Feranec J, Otahel J (2000) CORINE Land Cover Technical Guide Addendum 2000.Technical Report No 40. European Environment Agency, Copenhagen. Accessed on:

[9] http://ec.europa.eu/environment/water/flood_risk/index.htm

http://www.dmu.dk/fileadmin/Resources/DMU/Udgivelser/CLC2000/technical_guide_addenum.pdf

Bradshaw CJA, Sodhi NS, PehKSH, Brook BW (2007) Global evidence that deforestation amplifies flood risk and severity in the developing world. Global Change Biology 13(11): 2379-2395; doi: 10.1111/j.1365-2486.2007.01446.x

Burkhard B, de Groot R, Costanza R, Seppelt R, Jørgensen SE, Potschin M (2012a) Solutions for sustaining natural capital and ecosystem services. Ecological Indicators 21: 1-6

Burkhard B, Kroll F, Nedkov S, Müller F (2012b) Mapping ecosystem service supply, demand and budgets. Ecological Indicators 21: 17-29

Burkhard B, Kroll F, Müller F, Windhorst W (2009) Landscapes capacities to provide ecosystem services–a concept for land-cover based assessments. Landscape Online 15: 1-22

Crossman ND, Burkhard B, Nedkov S, Willemen L, Petz K, Palomo I, Drakou EG, Martin-Lopez B, McPhearson T, Boyanova K, Alkemade R, Egoh B, Dunbar MB, Maes J (2013) A blueprint for mapping and modeling ecosystem services. Ecosystem Services 4: 4-14

de Groot RS, Alkemade R, Braat L, Hein L, Willemen L (2010) Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. Ecological Complexity 7: 260-272

Guo Z, Gan Y (2002) Ecosystem function for water retention and forest ecosystem conservation in a watershed of the Yangtze River. Biodiversity and Conservation 11: 599-614

Kandziora M, Burkhard B, Müller F (2013a) Interaction of ecosystem properties, ecosystem integrity and ecosystem service indicators – A theoretical matrix exercise. Ecological Indicators 38: 54-78

Kandziora M, Burkhard B, Müller F (2013b) Mapping provisioning ecosystem services at the local scale using data of varying spatial and temporal resolution. Ecosystem Services 4: 47-59.

Laurance WF (2007) Environmental science: Forests and floods. Nature 449: 409-410; doi:10.1038/449409a.

Martinez-Harms MJ, Balvanera P (2012) Methods for mapping ecosystem service supply: a review. International Journal of Biodiversity Science, Ecosystem Services and Management 8: 17-25

Nedkov S, Nikolova M (2006) Modeling floods hazard in Yantra river basin. In: Proceedings from Balwois Conference, Ohrid, 23-26 May 2006.

Nedkov S, Burkhard B (2012) Flood regulating ecosystem services – mapping supply and demand, in the Etropole municipality, Bulgaria. Ecological Indicators 21:67-79

Nikolova M, Nedkov S, Semmens D, Iankov S (2007) Environmental quality and landscape risk assessment in Yantra River Basin. In: Petrosillo I, Müller F, Jones KB, Zurlini G, Krauze K, Victorov S, Li BL, Kepner WG (eds) Use of Landscape Sciences for the Assessment of Environmental Security. Springer, the Netherlands, p 202

Nikolova M, Nedkov S, Nikolov V (2008) Modeling local dimensions of the climate change in Etropolska Stara Planina. In: Papers from the International conference Global Environmental Change: Challenges to Science and Society in the South Eastern Europe, Sofia, 19-21 May 2008

Scott SN (2005) Implementing best management practices in hydrologic modeling using KINEROS and the Automated Geospatial Watershed Assessment (AGWA) Tool. Master's Thesis, School of Natural Resources, University of Arizona. 204p

Semmens D, Goodrich D, Uncrich C, Smith R, Woolliser D, Miller S (2005) KINEROS2 and the AGWA Modeling Framework. International G-WADI Modeling Workshop. http://gwadi.org/shortcourses/chapters/Semmens_L6.pdf

Stürck J, Poortinga A, Verburg PH (2014) Mapping ecosystem services: The supply and demand of flood regulation services in Europe. Ecological Indicators 38: 198-211

Vigerstol KL, Aukema JE (2011) A comparison of tools for modeling freshwater ecosystem services. Journal of Environmental Management 92:2403-2409

Part IV
Sensing Technologies and Their Integration with Maps

Sensing Technologies and their Integration with Maps: Mapping Landscape Heterogeneity by Satellite Imagery

Duccio Rocchini, Ana-Maria Olteanu-Raimond, Luca Delucchi, Sajid Pareeth, Markus Neteler, Harini Nagendra

Abstract Losses in biodiversity critically impact the ability of ecosystems to provide critical services ranging from carbon sequestration and food production to the maintenance of soil fertility. The maintenance of biodiversity is thus essential for human well-being and a sustainable future. Since landscape diversity often relates to species biodiversity, considering several ecological levels from species community diversity to genetic diversity, measuring landscape heterogeneity, is an efficient and relatively cheap way of providing biodiversity estimates over large geographical areas. In this study we will demonstrate the power of using remotely sensed data to estimate landscape heterogeneity and locate diversity hotspots, allowing effective management and conservation of the landscape.

1 Introduction

It is worth noting that the assessment of species diversity over relatively large areas is a challenging task. Compiling complete inventories has been hampered by the immense physical effort required for field estimates, and despite such effort, inaccurate estimates of diversity may result from changes in species composition through time.

Duccio Rocchini, Luca Delucchi, Sajid Pareeth and Markus Neteler
Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, Via E Mach 1, 38010 S. Michele all'Adige (TN), Italy, e-mail: duccio.rocchini@fmach.it, luca.delucchi@fmach.it, sajid.pareeth@fmach.it, markus.neteler@fmach.it

Ana-Maria Olteanu-Raimond
Laboratoire COGIT. Institut Géographique National, 2-4 avenue Pasteur 94165 Saint-Mandé Cedex, Paris, France e-mail: Ana-Maria.raimond@ign.fr

Harini Nagendra
Azim Premi University, PES Institute of Technology Campus, Pixel Park, B Block, Electronics City, Hosur Road Bangalore 560100, India e-mail: harini.nagendra@apu.edu.in

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_19,

Yet it is critically important to assess and monitor changes in species diversity for effective conservation. Losses in biodiversity critically impact the capability of ecosystems to provide critical services ranging from carbon sequestration and food production to the maintenance of soil fertility (Cardinale et al. 2012). Thus the maintenance of biodiversity is essential for human well-being and a sustainable future (Naeem et al. 2009; Nagendra et al. 2013).

New approaches have been proposed to overcome these issues, using landscape heterogeneity measured by the spatial variation of remotely sensed spectral signal as a proxy for species diversity (see Rocchini et al. 2010 for a review). The heterogeneity of the Earth's surface is closely related to physical and ecological diversity (Nagendra and Gadgil 1999; Gillespie et al. 2008). Since landscape diversity often relates to species biodiversity considering several ecological levels from species community diversity to genetic diversity, measuring landscape heterogeneity is an efficient and relatively cheap way of providing biodiversity estimates over large geographical areas. Depending on the study objectives, species diversity can also be modelled at appropriate scales in time and space. This is true in light of the Spectral Variation Hypothesis (Palmer et al. 2002; Rocchini 2007), which assumes that the higher the habitat heterogeneity, the higher will be the species diversity therein. Depending on the scale and the habitat being considered, the Spectral Variation Hypothesis can be expected to hold true in many cases.

The availability of satellite-derived data with high spatial (IKONOS, Orbview-3, BGIS-2000 (Balls Global Imaging System-2000), RapidEye) and spectral resolution (CHRIS (Compact High Resolution Imaging Spectrometer), Hyperion, GLI (Global Imager), MERIS (Medium Resolution Imaging Spectrometer), and MODIS (Moderate Resolution Imaging Spectrometer)) together with long term programmes such as the Landsat programme makes it feasible to study all terrestrial regions of the globe up to a resolution of few meters.

Free and Open Source tools (allowing the access to the source code and its peer-review) for assessing landscape heterogeneity at different spatial scales and in different environmental conditions (e.g. different habitats with divergent entropy gradients) are under development (e.g. Rocchini et al. 2013). Such tools may help to identify biodiversity hotspots from remotely-sensed and/or geographical and/or climatic data, which could help to focus field-based campaigns in a more efficient way in terms of time and costs, based on the application of best-fit-based parameters at appropriate spatial scales.

2 Information theory applied to the quantification of landscape heterogeneity

The most popular metrics for entropy measurement are derived from information theory, which measures the disorder contained in a system (Margalef 1958). In particular, consider a universe of entities u (e.g. pixels in a satellite image), each of which can be represented as a tuple z(u)—s(u), where z(u) = property of the uth entity related to its s(u) spatial component (Goodchild et al. 1999).

Most measures of spectral diversity that have been proposed thus far are based on either i) the Shannon entropy index (Shannon 1948; see also Bolliger 2005; Ricotta 2005) $H' = -\sum p \times \ln(p)$ with $0 \leq H' \leq \ln(N)$, where p is the relative abundance of each spectral reflectance value (DN) and N is the total number of possible values, or ii) reversed dominance, derived from the Simpson Dominance index $D = \sum p^2$ with $0 \leq D \leq 1$ (Simpson 1949), as $1-D$ (Simpson Diversity index). Both H' and $1-D$ will increase if the DN values are equally distributed with no DN value being dominant with respect to the others.

However, Nagendra (2002) and Rocchini and Neteler (2012) reported several problems with entropy based metrics, that is, a single index of diversity was not useful for either distinguishing different ecological situations, or for discerning differences in richness or relative abundance. This is because areas differing in richness or relative abundances of reflectance values (DNs) may share similar Shannon index values. Instead, coupling such entropy- or reversed-dominance-based metrics with indices taking into account evenness would dramatically increase the information content of such metrics. Among these, the mostly widely used index is the Pielou evenness index $J = \frac{-\sum p \times \ln(p)}{\ln(N)}$ (Pielou 1969) with $0 \leq J \leq 1$, which takes into account the maximum diversity given the same number of DNs and thus can be rewritten as $J = \frac{H'}{H'_{max}}$ (see also Ricotta and Avena 2003 for a review).

As demonstrated by Ricotta and Avena (2003), each index increases/decreases in a different manner depending on the relative array of abundances being considered. Hence, one could under- or over-estimate entropy/diversity depending on the metric being used. Diversity cannot be reduced to single index information, since one can never capture all aspects of diversity in a single statistic (Gorelick 2006). As an example, Nagendra (2002), dealing with the Shannon H' index and the Simpson diversity index $1-D$, reports the case of discordant diversity patterns obtained by considering different indices. Such information may remain hidden if only one index is considered. Thus, following O'Neill et al. (1988) in a pioneer study on the landscape indices, a restricted set of non-redundant indices could reach significant aspects on the spatial patterns.

In this view, generalised entropy based algorithms for entropy calculation could provide a solution to such problems since they encompass a continuum of diversity measures by varying one of few parameters in their formula.

As an example, Rényi (1970) proposed a generalised entropy:

$$H_\alpha = \frac{1}{1-\alpha} \ln(\sum p^\alpha) \qquad (1)$$

where p=relative abundance of each spectral reflectance value (DN).

Such measure is extremely flexible and powerful since many popular diversity indices are simply special cases of H_α.

For instance, for α=0, $H_0 = \ln(N)$ namely the logarithm of richness (N=number of DN values) or the maximum Shannon entropy index (H_{max}) which is used as the denominator of the Pielou index, while for α=2, $H_2 = \ln \frac{1}{D}$ where D is the Simpson Dominance index. For α=1 the Rényi entropy is not defined and its derivation, H_1=Shannon's entropy H', is based on l'Hôpital's rule of calculus (Ricotta 2005).

While traditional metrics supply point descriptions of diversity, in Rényi's framework there is a continuum of possible diversity measures, which differ in their sensitivity to rare and abundant DNs, becoming increasingly regulated by the commonest DNs when increasing the values of α. This is why Rényi generalised entropy has been referred to as a "continuum of diversity measures" (Ricotta et al. 2003).

Implementing such algorithms into Open Source Software will help researchers to freely develop measures of landscape heterogeneity in an Open Source code space.

The aim of this study is to demonstrate the power of using remotely sensed data to estimate landscape heterogeneity and locate diversity hotspots over space, allowing effective management and conservation of the landscape. We will rely on the Rényi generalised entropy, based on the Free and Open Source Software GRASS GIS.

3 Open Source software philosophy for the calculation of landscape diversity metrics

The idea of Free and Open Source (FOSS) software has been around for almost as long as software has been developed (Neteler and Mitasova 2008).

The famous "four freedoms" paradigm, developed by Richard Stallman (1985) in his seminal work, proclaims i) the freedom to run the program for any purpose, ii) the freedom to study how the program works and adapt it to one's own needs, iii) the freedom to redistribute copies, and iv) the freedom to improve the program and release such improvements to the public. This guarantees that the whole community benefits from software development (also see Fogel 2009).

With the aim of calculating landscape metrics in GIS to ensure robust analysis output, particularly where complex algorithms are concerned (Neteler and Mitasova 2008), full access to the source code is crucial. There are well-known examples of FOSS in research fields such as statistics (e.g. R Language and Environment for Statistical Computing, R Development Core Team, 2013), while GIS scientists and more generally landscape ecologists may benefit from the powerful GIS named GRASS (Geographical Resources Analysis Support System, http://grass.osgeo.org, see Neteler et al. 2012), which includes more than 350 modules for managing and analysing geographical data. GRASS was originally created in 1982 by the U.S. Army Construction Engineering Research Laboratories, and is now one of the cutting-edge projects of the Open Source Geospatial Foundation (OSGeo, founded in 2006). Quoting Neteler and Mitasova (2008):

> The key development in recent GRASS history was the adoption of GNU GPL (General Public License, see http://www.gnu.org) in 1999. With this, GRASS embraced the Open Source philosophy, well known from the GNU/Linux development model, which stimulated its wide acceptance.

Adoption of the FOSS license changed the development process of GRASS with contributions to the source code becoming decentralised. The legal statements declared in the GPL are based on the aforementioned "four freedoms" paradigm (Stallman 1985; 1997) and allow the user to use the software's full range of capabilities, and to distribute, study and improve it.

Figure 19. 1 from Rocchini et al. (2013) represents an example of Rényi calculation (with $\alpha = 2$, relying on the *r.diversity* function of GRASS GIS (available at http://grasswiki.osgeo.org/wiki/AddOns/GRASS6), and the comparison with the Shannon, Simpson and Pielou diversity indices. Notice that the Simpson diversity index results in an emphasissmall differences in low-diversity areas (e.g. in homogeneous zones) since its formula contains a squared p, while logarithm-based indices (Shannon entropy, Pielou evenness, Rényi generalised entropy) enhance differences in sites with higher evenness (Nagendra, 2002).

4 A case study: mapping spatial heterogeneity in the Tadoba Andhari Tiger Reserve, India

The Tadoba Andhari Tiger Reserve (TATR) is a national park and wildlife sanctuary located in central India, in the eastern part of Maharashtra state. The protected area extends over 625 square kilometres, covering a landscape that is largely a matrix of dry tropical forests, interspersed with some grasslands, water bodies and a few small patches of riparian forest alongside streams. The park is drained by two main rivers. The southern section of the park is flat, giving way to gradually undulating topography as one moves northwards.

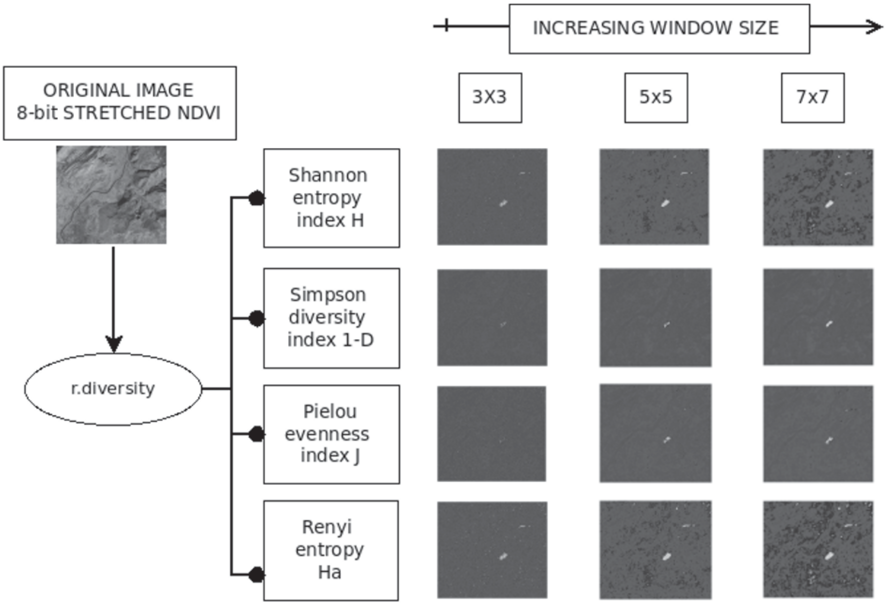


Fig. 19. 1 An example of the calculation of the Rényi generalised entropy with $\alpha = 2$ and the comparison with the Shannon, Simpson and Pielou indices, based on the *r.diversity* function of GRASS GIS applied to a Landsat Normalized Difference Vegetation Index (NDVI) map. Reproduced from Rocchini et al. (2013) under license permission number 3331820502202 from Elsevier Ltd. Refer to the main text for additional information.

There is a well-developed road network in the northeastern part of the reserve, which provides access to the forest for grazing and biomass extraction. To the north, south and east, the TATR has some protection from surrounding State controlled Reserve Forest and Protected Forest areas. Six villages are located within the boundaries of the park (one village has since been relocated outside the park), and there are two villages located on the periphery. In addition to the six interior villages, several other villages and communities access resources from the park. Nagendra et al. (2010a) individualised three main zones of human impact/pressure as reported in Figure 19. 2. These villages fulfil a large part of their fuel, fodder, timber and non-timber forest requirements from the park. The TATR also experiences substantial seasonal use from migrant herders, and is frequented by timber, bamboo and wildlife poachers. Thus, despite being located within a protected area, this dry tropical forest habitat is also subject to human disturbance due to grazing, fire and biomass extraction Figure 19. 2.

A Landsat ETM+ image, acquired on 29 October 2001 (path 144, row 046, spatial resolution 28.5 meters, band from 1 to 5 and 7. see Rocchini et al. 2009 and http://glcfapp.glcf.umd.edu:8080/esdi/ftp?id=268517)
covering the whole study area, was downloaded from the Global Land Cover Facil-

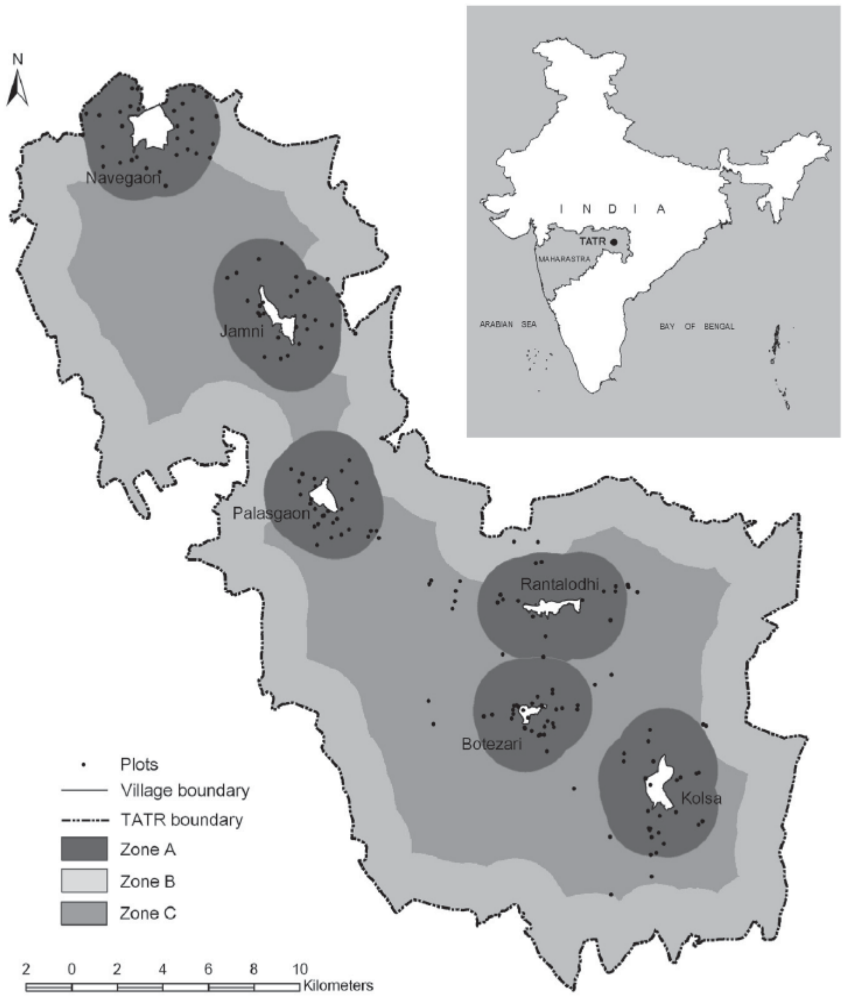


Fig. 19. 2 Map of the Tadoba Andhari Tiger Reserve (TATR, India), showing the villages in the park and three zones of human impact, from high (zone A, villages' surroundings) to low (zone C, dry tropical forest and grasslands) human disturbance, as defined in Nagendra et al. (2010a). The figure is reproduced from Nagendra et al. (2010a) under license permission number 3331821040245 from Elsevier Ltd.

ity site hosted by the University of Maryland (www.glcfapp.umiacs.umd.edu, Tucker et al. 2004 for major details).

A Principal Component Analysis (PCA) was performed on the image and the first component explaining 49.57% of the total variance was retained and rescaled to 16-bit to calculate the landscape heterogeneity by the *r.diversity* function in GRASS GIS.

Figure 19. 3 shows the results attained when relying on Rényi diversity with α=0 (related to richness) and α=2 (related to evenness). While Rényi diversity based on pure richness immediately saturated towards highest values (pixels being different from each other in a window of 3x3 pixels), this effect was not achieved when considerig evenness (Figure 19. 3) where the contrast between diverse and non-diverse areas is higher. It is worth noting that the highest diversity, considering overall evenness (α=2, i.e. contrast in the spectral signal) was found not only in the dry natural forest of the study area but also in the villages, as previously postulated by Nagendra et al. (2010a). Looking as an example at the location of the Navegaon village (Figure 19. 2, northern part) in the Rényi diversity index images (Figure 19. 3), this is one of the areas with highest landscape diversity over the whole area.

5 Discussion

The Tadoba Andhari Tiger Reserve (TATR) is an important tiger reserve in India, containing a large contiguous habitat of dry tropical forest that is very important for tigers as well as other large wildlife that is characteristic of the central Indian landscape. The landscape is also a very important source of human livelihood, with high densities of forest-dependent communities living in this region who harvest a wide range of forest products including timber, bamboo and non-timber forest products such as medicinal herbs (Nagendra et al. 2006). Hence, in addition to its biological significance, the maintenance of biological diversity has great social importance for livelihood sustainability. The forest dependent tribal populations living within the park are extremely dependent on the biodiversity in this forest for their daily livelihood, using as many as 19 different species of trees for timber and an even larger number of species for medicinal use through personal consumption, as well as for sale (Nagendra et al., 2010a). Depletion of biological diversity will create extreme problems for these communities.

Mapping and monitoring of biodiversity and forest heterogeneity in the TATR is thus very important for a better understanding of human impacts on ecology, given the importance of this large heterogeneous landscape for biodiversity. Such studies also provide a deeper understanding of human impacts on dry tropical forest, which represents a habitat type that is very important for biodiversity because of its biological richness as well as its vulnerability to human impact, yet which remains little studied in comparison to moist tropical and dry temperate forests (Feeley et al. 2005). As this research demonstrates, the villages in the interior of the park contain significant amounts of spatial heterogeneity which is believed to relate to biological diversity (Nagendra et al. 2010b). In particular, Navalgaon, a village that is located at the park gate with high levels of disturbance but which also provides access to the less disturbed park interior, has a significant diversity of habitats which relates to the intermediate disturbance hypothesis (Paine and Vadas 1969; Grime 1973; see also Catford et al. 2012 and references therein), which indicates that areas of intermediate disturbance are likely to have maximum diversity. This research thus points

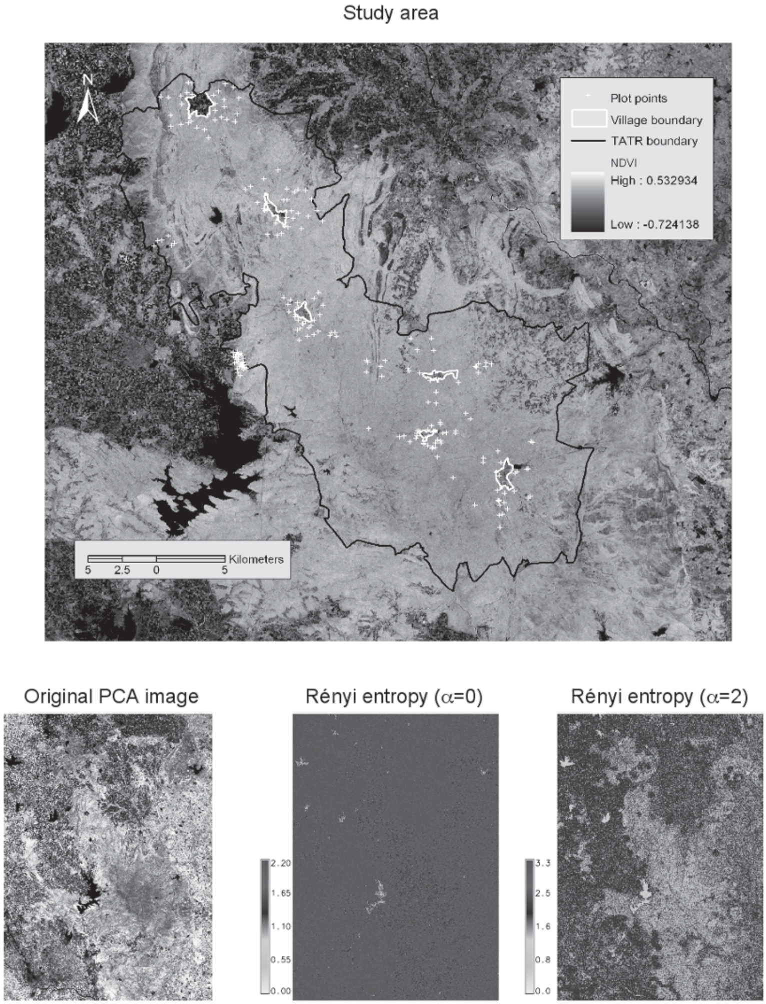


Fig. 19. 3 The Tadoba Andhari Tiger Reserve (TATR, NDVI image on top from Nagendra et al. 2010b, with village borders in white) and the calculation of the Rényi generalised entropy, measured from the first principal component of a Landsat ETM+. Rényi diversity based on pure richness immediately saturated towards highest values; this effect was not achieved when considering evenness where the contrast among diverse and non-diverse areas is higher. It is worth noting that the highest diversity was found not only in the dry natural forest of the study area but also in the villages. Refer to the main text for additional information.

to a need to go beyond the treatment of parks as monoliths requiring a standardized approach for management band monitoring across all areas(Nagendra et al. 2010a), towards a more sophisticated use of mapping approaches (e.g. Nagendra et al. 2013) to provide input for spatially directed monitoring and management based on the differences in heterogeneity and diversity across different areas of the park.

Despite efforts to demonstrate the relationship between remotely sensed- and species-diversity, there are still no useful tools available for managers and administrators responsible for environmental policy and landscape diversity, even though the maintenance of diversity is often a stated objective of these decision makers (Nagendra 2002). Strictly speaking, it is becoming increasingly important to develop means for rapidly and objectively forecasting species diversity in order to assess, with limited resources, the impacts of anthropogenic and natural disturbances on biodiversity.

Due to the difficulties of field-based data collection at wider spatial scales, we have demonstrated in this study how the use of remote sensing for estimating environmental heterogeneity as a proxy of diversity at different spatial scales represents a powerful tool since it allows an a-priori estimate of potential hotspots of diversity allowing an effective management and conservation of the landscape.

Acknowledgements D.R. is partially supported by the EU BON (Building the European Biodiversity Observation Network) project, funded by the European Union under the 7th Framework programme, Contract No. 308454. D.R. and A.-M. O.-R. are partially supported by the ICT COST Action TD1202 “Mapping and the citizen sensor”, funded by the European Commission.

Appendix 1:

Code used to calculate Rényi entropy in GRASS GIS 7.0

Import landsat bands

```
# freely available from http://glcfapp.glcf.umd.edu:8080/esdi/
# path of the used image: 144
# row of the used image: 046
# acquisition date: October 29th 2011
r.in.gdal in=/path/p144r046_7dt20011029.SR.b01.tif out=B1
r.in.gdal in=/path/p144r046_7dt20011029.SR.b02.tif out=B2
r.in.gdal in=/path/p144r046_7dt20011029.SR.b03.tif out=B3
r.in.gdal in=/path/p144r046_7dt20011029.SR.b04.tif out=B4
r.in.gdal in=/path/p144r046_7dt20011029.SR.b05.tif out=B5
r.in.gdal in=/path/p144r046_7dt20011029.SR.b07.tif out=B7
```

Set the region of analysis (optional)

```
g.region=B7 # all bands are overlapping each other
# image information
```

```
# projection: 1 (UTM)
# zone: 44
# datum: wgs84
# ellipsoid: wgs84
# north: 2292630
# south: 2194170
# west: 293250
# east: 362610
# nsres: 30
# ewres: 30
# rows: 3282
# cols: 2312
# cells: 7587984
```

Perform PCA to extract one single band with the highest information content

```
i.pca input=B1,B2,B3,B4,B5,B7 output_prefix=pc rescale=0,255  then store pc1
```

Perform Rényi entropy calculation

```
# code to produce entropy maps of Figure 19. 3
r.diversity input=pc.1 prefix=pc.1 size=3 method=renyi alpha=0
r.diversity input=pc.1 prefix=pc.1 size=3 method=renyi alpha=2
# the size equals the moving window size of analysis, in this case a 3x3
# alpha parameter is related to the Rényi formula (see the main text)
```

References

Bolliger J (2005) Simulating complex landscapes with a generic model: Sensitivity to qualitative and quantitative classifications. Ecol Complex 2:131-149

Cardinale BJ, Srivastava DS, Duffy JE, Wright JP, Downing AL, Sankaran M, Jouseau C (2005) Biodiversity loss and its impact on humanity. Nature 486: 59-67

Catford JA, Daehler CC, Murphy HT, Sheppard AW, Hardesty BD, Westcott DA, Rejmnek M, Bellingham PJ, Pergl J, Horvitz CC, Hulme PE (2012) The intermediate disturbance hypothesis and plant invasions: Implications for species richness and management. Perspectives Plant Ecol Evol Systematics 14, 231-241

Feeley KJ, Gillespie TW, Terborgh JW (2005) The utility of spectral indices from Landsat ETM+ for measuring the structure and composition of tropical dry forests. Biotropica 37:508-519

Fogel K (2009) Producing open source software: how to run a successful free software project. Creative Commons Attribution-ShareAlike (3.0) license. O'Reilly, Sebastopol, USA

Gillespie TW, Foody GM, Rocchini D, Giorgi AP, Saatchi S (2008) Measuring and modelling biodiversity from space. Progr Phys Geogr 32:203-221

Goodchild MF, Egenhofer MJ, Kemp KK, Mark DM, Sheppard ES (1999) Introduction to the Varenius project. Int J Geogr Inf Sci 13:731-745

Gorelick R (2006) Combining richness and abundance into a single diversity index using matrix analogues of Shannon's and Simpson's indices. Ecography 29: 525-530

Grime JP (1973) Competitive exclusion in herbaceous vegetation. Nature 242:344-347

Margalef R (1958) Information theory in ecology. Gen. Syst. 3:36-71

Naeem S, Bunker DE, Hector A, Loreau M, Perrings C (2009) Biodiversity, Ecosystem Functioning, & Human Wellbeing. An Ecological and Economic Perspective. Oxford University Press, Oxford, UK

Nagendra H, Gadgil M (1999) Biodiversity assessment at multiple scales: linking remotely sensed data with field information. Proc Natl Acad Sci USA 96:9154-9158

Nagendra H, Lucas R, Honrado JP, Jongman RHG, Tarantino C, Adamo M, Mairota P (2013) Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity and threats. Ecol Indic 33:45-59

Nagendra H (2002) Opposite trends in response for the Shannon and Simpson indices of landscape diversity. Appl Geogr 22, 175-186

Nagendra H, Pareeth S, Ghate R (2006) People within parks: forest villages, land-cover change and landscape fragmentation in the Tadoba-Andhari Tiger Reserve, India. Appl Geogr 26:96-112

Nagendra H, Reyers B, Lavorel S (2013) Impacts of land change on biodiversity: making the link to ecosystem services. Current Opinions Environ. Sustainability 5:503-508

Nagendra H, Rocchini D, Ghate R (2010a) Beyond parks as monoliths: Spatially differentiating park-people relationships in the Tadoba Andhari Tiger Reserve in India. Biol Conserv 143:2900-2908

Nagendra H, Rocchini D, Ghate R, Sharma B, Pareeth S (2010b) Assessing plant diversity in a dry tropical forest: Comparing the utility of Landsat and IKONOS satellite images. Remote Sens 2:478-496

Neteler M, Bowman MH, Landa M, Metz M (2012) GRASS GIS: a multi-purpose Open Source GIS. Environ Model Software 31:124-130

Neteler M, Mitasova H (2008) Open Source GIS: A GRASS GIS Approach, The International Series in Engineering and Computer ScienceThird Edition. Springer, New York, USA

O'Neill RV, Krummel JR, Gardner RH, Sugihara G, Jackson B, DeAngelis DL, Milne BT, Turner MG, Zygmunt B, Christensen SW, Dale VH, Graham RL (1988) Indices of landscape pattern. Landscape Ecol 1:153-162

Paine RT, Vadas RL (1969) The effects of grazing by sea urchins, *Strongylocentrotus* spp., on benthic algal populations. Limnol Oceanogr 14:710-719

Palmer MW, Earls P, Hoagland BW, White PS, Wohlgemuth T (2002) Quantitative tools for perfecting species lists. Environmetrics 13:121-137

Pielou, EC (1969) An Introduction to Mathematical Ecology. Wiley, New York, USA

R Development Core Team (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. (

http://www.R-project.org,)

Rényi A (1970) Probability Theory. North Holland Publishing Company, Amsterdam, The Netherlands

Ricotta C (2005) On possible measures for evaluating the degree of uncertainty of fuzzy thematic maps. Int J Remote Sens 26:5573-5583

Ricotta C, Avena G (2003) On the relationship between Pielou's evenness and landscape dominance within the context of Hill's diversity profiles. Ecol Indic 2:361-365

Ricotta C, Corona P, Marchetti M, Chirici G, Innamorati S (2003) LaDy: software for assessing local landscape diversity profiles of raster land cover maps using geographic windows. Environ Model Software 18:373-378

Rocchini D (2007) Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. Remote Sens Environ 111: 423-434

Rocchini D, Balkenhol N, Carter GA, Foody GM, Gillespie TW, He KS, Kark S, Levin N, Lucas K, Luoto M, Nagendra H, Oldeland J, Ricotta C, Southworth J, Neteler M (2010) Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. Ecol Inform 5:318-329

Rocchini D, Delucchi L, Bacaro G, Cavallini P, Feilhauer H, Foody GM, He KS, Nagendra H, Porta C, Ricotta C, Schmidtlein S, Spano LD, Wegmann M, Neteler M (2013) Calculating landscape diversity with information-theory based indices: A GRASS GIS solution. Ecol Inform 17:82-93

Rocchini D, Nagendra H, Ghate R, Cade BS (2009) Spectral distance decay: assessing species beta-diversity by quantile regression. Photogramm Engin Remote Sens 75:1225-1230

Rocchini D, Neteler M (2012) Spectral rank-abundance for measuring landscape diversity. Int J Remote Sens. 33:4458-4470

Shannon C (1948) A mathematical theory of communication. Bell Syst Tech J 27:379-423

Simpson EH (1949) Measurement of diversity. Nature 163:688

Stallman R (1985) The GNU Manifesto Available at:
http://www.gnu.org/gnu/manifesto.html

Stallman R (1997) The GNU Manifesto, In: Ermann MD, Williams MB, Shauf MS (Eds.), Computers, Ethics and Society, Second Edition. Oxford University Press, Oxford, pp. 229-239

Tucker CJ, Grant DM, Dykstra JD (2004) NASA's global orthorectified Landsat data set. Photogramm Engin Remote Sens 70:313-322

Analysis of Image Transformation and Land Use/Land Cover for Temperature Trends on Landsat Imagery

Iswari Nur Hidayati, Khalifah Insan Nur Rahmi, Gerry Kristian, Vidya Nahdhiyatul Fikiyah, Diyah Puspitaningrum

Abstract The global climate change phenomenon has become a big issue, one aspect of which is change in temperature (temperature trend) in urban areas. Temperature trend can be investigated using a remote sensing data approach through multiple extractions based on transformations such as NDVI and NDBI. This transformation could represent land cover types quantitatively so the relationship between different indices, such as NDVI, NDBI, and the temperature can be seen. Processing shows that the distribution and trend of temperatures in a Landsat 5 TM thermal band for periods 1992 and 2009 has risen by an average temperature of 0.99˚C, an indicator of the temperature trend in Yogyakarta city. The results of the relationship between brightness temperature and land use/cover pattern (LUCP) is seen in the results of the transformation of NDVI and NDBI correlated with brightness temperature values. The correlation value for temperature and those two parameters suggests that the NDBI transformation is the most dominant factor affecting the increase in temperature, with a regression coefficient reaching 0.5184, and that larger built up areas will create hotter temperatures.

1 Introduction

Urbanisation has increased in recent years, including in regions of Indonesia, and this will affect population growth in urban areas and this will mean changes in the way land is used in built-up areas. The changes will impact the average air temperature in the urban area; where vegetation is replaced by the built-up area there will be a greater contrast between urban and rural area of surface radiance and air temperature after this change. (Weng 2001, Hawkins et al. 2004, Lin et al. 2011).

Studies of the land surface commonly include large cities and other urban areas, but actually the change of temperature in urban areas occurs not only in cities but also in small towns. Land use change in the big cities is more frequent, however and the differences in air temperature are greater than in small towns. All cities of any size form their own distinct climate with the regional macroclimate in which the city located, although the characteristics of the urban microclimate depends on the larger climate conditions (Kopec 1970).

Cartographyy and Remote Sensing, Faculty of Geography,
Universitas Gadjah Mada, Bulaksumur, Yogyakarta, Indonesia
Phone (+62274) 902340
Email: iswari@geo.ugm.ac.id

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_20,

Temperature trend has been a concern for more than 40 years. Deosthali (2000) found that at the night, the city core appears in the night as heat and humidity island and in a daylight as heat and dry islands. Saaroni et al. (2000) created a new method for monitoring the temperature trend combinations of different scales, which allows for the assessment of a broad range of urban spatial temperature trends and thermal characteristics. Rao (1972) was the first to demonstrate that urban areas can be identified from the analysis of thermal infrared data using satellite remote sensing imagery. Studies of temperature trends using remote sensing are also conducted by Betchel et al.(2012), to monitor the land surface temperature and results have shown that the increase of urbanization has caused changes in the heat balance in built-up areas, where the urban centre has a higher temperature that the rural surroundings. In this case, remote sensing provides the advantages of multitemporal data. There have been studies of the temperature trend with the approach of land use using Land Surface Temperature (LST) from NOAA AVHRR imagery in urban areas (Gallo and Owen 1998, Streutker 2002, 2003). More recently, Landsat Thermatic Mapper (TM) and Enhanced Thermatic Mapper Plus (ETM +) imagery and thermal infrared (TIR) have also been used for research into temperature change (Chen et al. 2002, Weng 2001). This means that remote sensing can be a supporting instrument to derive environmental characteristics as well as providing many options to determine objects.

This qualitative study of the relationship between land use/cover patterns (LUCP) and LST will assist in the land use planning process. It is known that the vegetation index derived from remote sensing can be used in the assessment of vegetation cover in both a qualitative and a quantitative way (Tian and Min 1998). The relationship between the vegetation indexes of various vegetation cover have been established using regression analysis (Purevdorj et al. 1998), as for the Ratio Vegetation Index (RVI), Normalised Difference Vegetation Index (NDVI), Difference Vegetation Index (DVI) and Perpendicular Vegetation Index (PVI). NDVI has been used to estimate the productivity of vegetation and rainfall in semi-arid regions (Chen et al. 2005, Wang et al. 2004). Zha et al. (2003) developed the Normalised Difference Built-up Index (NDBI) to identify urban areas and built- up areas. It is possible that the use of NDVI and NDBI could represent quantitative land cover types so that the relationship between different indexes, such as NDVI and NDBI, and the temperature can be included in the study of temperature trend.

From this background it can be seen that several transformations, such as those of the NDVI and NDBI, can be used to extract information about land use or other data to obtain information that can be used as in analysis of the temperature trend and urban heat islands (Cox et al. 2005, Liu and Zhang 2011). The extraction process uses multitemporal data so as to reveal the trend in temperatures. One of the satelite that can provide thermal data is Landsat, which has a spatial resolution of 30m for Landsat 5 TM. The thermal data can be processed to produce a Land Surface Temperature (LST), using radiance value calculations, to generate surface temperature values (Zhao-ming et al. 2006).

The objectives of this research are to determine the distribution and development in the temperature trend using Landsat TM/ETM+ in thermal band for the periods 1992 and 2009, to analyse the relationship between brightness temperature and land use/land cover pattern (LUCP) in multitemporal conditions. This study therefore focused on the distribution of temperature trends in correlation with land use using remote sensing data.

2 Data and Methods

This research employs a quantitative methodology, using remote sensing as the data source for processing. Data from the thermal sensor of satellite imagery was used to detect the surface temperature and the multispectral image data to identify the Normalised Difference Built Up Index (NDBI), vegetation density index (NDVI), and land use/land cover. The relationship between NDBI, NDVI, and land use/land cover and surface temperature is determined after processing and field checks on the data.

2.1 Study Area

The study area is the city of Yogyakarta, in Daerah Istimewa Yogyakarta Province, Indonesia. It is located at approximately the centre of the province, in the southern part of Java island. It is the largest city of the province with a population in 2012 of 394,012 persons (BPS 2012).

2.2 Data Sources

Analysis of the temperature trend uses the Landsat 5 TM full band recordings of 1992 and 2009. The system on Landsat 5 is designed to collect the reflected energy carried by bands 1-5 and 7 (6 bands) and the energy carried by band 6 (1 band). Landsat sensors convert the reflected solar energy into radiance units. Radiance is defined as the radiant flux per unit solid angle leaving an extended source in a given direction per unit projected surface area in a particular direction. This radiance is closely related to the brightness in the specific direction of the sensor. Radiance is measured by the sensor and related to reflection. Radiance values are then quantified into brightness values images and stored in a digital format (Sutanto 1994b).

A topographic map of Yogyakarta at the scale of 1: 25.000 and Quickbird imagery of Yogyakarta was used to derive the data of land use/land cover patterns, which is then linked to the trend temperature. The fieldwork to conduct the ceck existing of temperature and land use that will be the accuracy test of interpretation derived from remote sensing data, and to collect the field data uses surveying tools such as GPS, infrared thermometers, digital cameras, check lists, and gauges.

2.3 Processing the brightness temperature from Landsat 5 TM

Thermal data obtained from the imagery is initially a digital value (Digital Number/ DN), so the value should be converted in radiance. Chen (2002) conducted a study to obtain the data brightness temperature using two steps, by converting the digital number of band 6 to radiance values using the following formula :

$$R_{TM6} = \frac{V}{255}(R_{max} - R_{min}) + R_{min} \qquad (1)$$

Where : V = pixel value
R_{max} = 1.896 ($mW*cm^{-2}*sr^{-1}$)
R_{min} = 0.1534 ($mW*cm^{-2}*sr^{-1}$)

The value is then converted to Kelvin temperature using the following formula:

$$T = \frac{K1}{\ln\left(\frac{\frac{K2}{R_{TM6}}}{b}\right)+1} \qquad (2)$$

Where : K1 = 1260.56 K
K2 = 60.766 ($mW*cm^{-2}*sr^{-1}\mu m^{-1}$)
b = range of effective spectral value (1.239 μm)

There were changes in the formula to change the value of the digital number (DN) to the surface temperature value due to the mismatch of the formula used for the image. Temperature values in 1992 and 2009 were obtained from the Landsat 5 image. Based on the research of Chander and Markham (2003) the transformation to convert the digital number values (DN) into radiance values is based on a formula as follows:

$$L_{\tau} = \left(\frac{Lmax_{\tau} - Lmin_{\tau}}{Q_{cal\,max}}\right) Q_{cal} + Lmin_{\tau} \qquad (3)$$

Where: L_{τ} = spectral radiance (watts/(m^2*ster*μm)
$Lmax_{\tau}$ = spectral radiance that is scaled to Qcal-max(watts/(m^2*ster*μm)

$Lmin_{\tau}$	= spectral radiance that is scaled to Qcalmin (watts/(m^2*ster*µm)
$Q_{cal\ max}$	= maximum quantized calibrated pixel (DN=255)
Q_{cal}	= maximum quantized calibrated pixel (DN=0)

Converting radiance value to surface temperature value based on a formula is as follows:

$$T = \frac{K2}{\ln\left(\frac{K1}{L\tau}+1\right)} \quad (4)$$

Where:

	T	= brightness temperature (K);
	K1	= 666.09 (watts/(m^2*ster*µm)
	K2	= 1282.71 (Kelvin)
	Lλ	= spectral radiance (watts/(m^2*ster*µm)

2.4 Transformation Vegetation Index

A vegetation index is an algorithm applied to an image (multi-band), to highlight aspects of vegetation density or other aspects related to the density, for example biomass, Leaf Area Index (LAI), chlorophyll concentration, and so on (Danoedoro 1996). In vegetated areas, the infrared part of the spectrum is reflected by the leaves and the reflection is received by the sensor, and on the red spectrum it is absorbed by chlorophyll, thereby reducing the reflection of the red light detected by the sensor. The reflectance difference can be used to evaluate the presence of vegetation (Lillesand and Kiefer 1997).

2.4.1 Transformation of NDVI from Landsat 5 TM

This research used NDVI (Normalized Difference Vegetation Index), because it is sufficiently established compared to other vegetation transformations. In addition, this transformation has been widely used in remote sensing application sharing. The equation used is (Danoedoro 2012):

$$NDVI = \frac{\rho(band\ 4)-\rho(band3)}{\rho(band4)+\rho(band3)} \quad (5)$$

2.5 Transformation Index of Built-Up Area

NDBI is also called the Normalised Difference Built-up (Barren) Index. The built-up index is an algorithm to show the density of a built-up area/barren soil (Friedl and Broadley 1997). NDBI is very sensitive to built-up or open areas. This algorithm uses band 5 of Landsat 7 ETM+ to highlight the moisture content in a variety of land use, or to differentiate land and buildings from the other land use. Band 4 is very sensitive and has a high reflectance of vegetation, while the reflectance of open land and built land is very low.

2.5.1 Transformation of NDBI from Landsat 5 TM

Built-up index transformation in the present study uses the NDBI equation. This transformation was chosen because it is one of transformations that has been frequently used to determine the built up index. The equation for the Landsat 5 TM is as follows (Danoedoro, 2012):

$$NDBI = \frac{d(band5)-d(band4)}{d(band5)+d(band5)} \qquad (6)$$

Sample Selection

The sample selection is considered by the mapping unit. The mapping unit is composed of temperature maps, NDVI maps, NDBI maps, and land cover/land use map. At least 1 sample was taken from every mapping unit. The preparation of NDVI maps, NDBI maps, and temperature maps was based on the distribution of pixels and the specified object classes.

Accuracy Assessment

Interpretation test activities were performed to determine the accuracy of the data generated in the laboratory activities. The interpretation test for land cover/land use is measured quantitatively and expressed in percentages, with the following formula (Sutanto 1994a):

$$Level\ of\ accuracy = \frac{number\ of\ appropriate\ sample}{number\ of\ total\ sample} x\ 100\% \qquad (7)$$

Field Data Collection

Surface temperature data was obtained through direct measurements in the field using an infrared thermometer, which is a tool that is able to measure the temperature directly on the surface (Daldjoeni 1986). Temperature measurements were performed on each sample of the mapping unit, so that the variation could be obtained from a variety of combined elements.

Built-up area density and vegetation density data was collected from the fieldwork. The field work was conducted to check the accuracy of the transformation of NDVI and NDBI, with the help of high-resolution images to show both the vegetation density and built-up area.

Calculation Statistic Data

A correlation test was used to determine the association degree and influence degree.

3 Results and Discussion

3.1 Imagery radiometric correction

Radiometric correction of Landsat 5 TM images used in this study transformed the digital numbers of band 6 into radian values based on certain algorithms. The value of Rmax and Rmin that was used as algorithms is 1.896 mW cm-2 sr-1; and 0.1534 mW cm-2 sr-1. The results indicated the changed pixel values of the image that had not been corrected by the image have now been corrected. A histogram of the results of the radiometric correction of Landsat 5 TM image of 1992 showed the minimum and maximum pixel values of 8.436, and 10.042 which originally amounted to 131, and 160 (Annex 1.1). The distribution of pixel values before and after radiometric correction of Landsat 5 TM images in 1992 and 2009 can also be seen in Table 20.1.

Table 20.1 Pixel value distribution of Landsat 5 TM images before and after radiometric correction

Year	Pixel value			
	Radiometric correction	Min	Max	Mean
1992	Before	123	162	145,574

	After	8,436	10,042	9,123
2009	Before	123	162	145,574
	After	7,993	10,153	9,244

Determination of the value of brightness temperature utilised band 6 on the Landsat 5 TM image recording in 1992 and 2009, which is a thermal infrared band with wavelength of 10.40 to 12.50 µm. The result of images processing then can be seen in Figure 20.1.

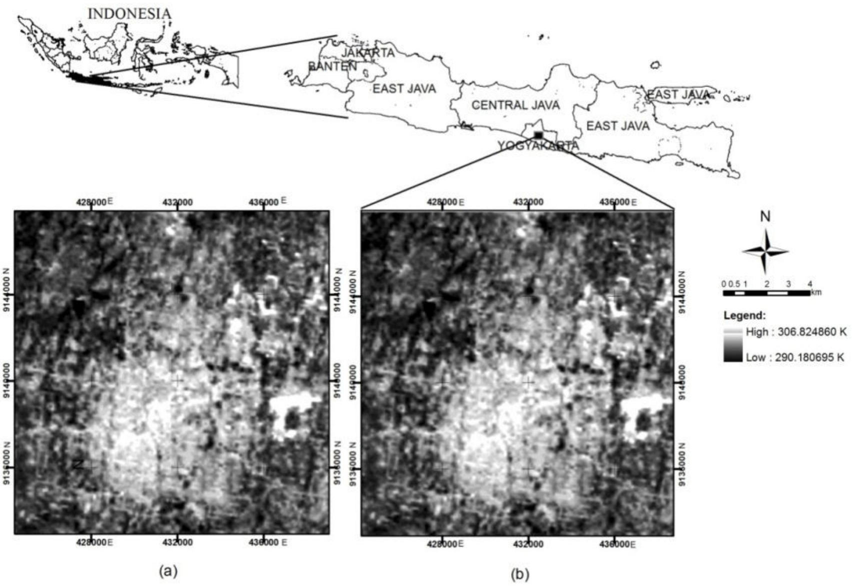


Fig. 20.1 Thermal image processing results in (a) 1992 and (b) 2009

These results (Figure 20.1) show the magnitude of brightness temperature at each pixel, taking a random sample of 100 research sites to obtain the magnitude of the temperature of each land use. The results indicated the presence of temperature changes in each land use from 1992 to 2009, and temperature changes of 0.99^{o} C. As an example in the form of land use was built-up area in 1992 originally had a temperature of 27.575^{o} C, up about 1.2^{o} C to about 28.775^{o} C.

3.2 Land Use Classification

Interpretation and analysis of remote sensing imagery involves the identification of various targets in an image in order to extract useful information. Remote sensing data is used for supervised classification for creating land use/land cover data. The advantage is that it can also analyse multi-temporal data (time series). Remote sensing data can compared land use in 1992 and 2009. The land use changes in 1992 and 2009 can become references for making maps of temperature, because land use change affected the changes of temperature especially in urban areas.

The land use classification in this study used the multispectral classification algorithm, which is based on the value of the spectral classification of each pixel simultaneously on multiple bands. The supervised classification with maximum likelihood algorithm was used for land use classification. The results indicated that there are a number of land uses ranging from built-up land, roads, bare land, high vegetation density, and low vegetation density (Figure 20.2).

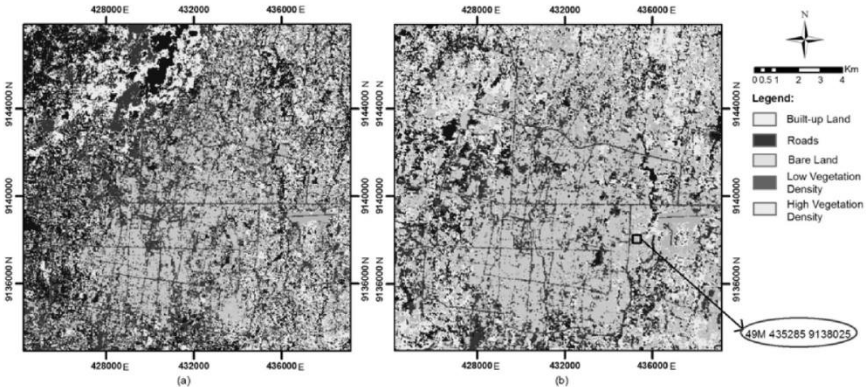


Fig. 20.2 Results of the multispectral classification, maximum likelihood algorithm (a) 1992, (b) 2009

The classification results of land use in 1992 (look coordinate sample) showed a high density of vegetation, but 17 years later, in 2009, indicate a change in land use to that of a built-up area. Land use change is justified by the presence of field activity which indicates that the site is a settlement. Based on 100 randomly selected samples, it could be seen that the change of land use trends towards a built-up area, whether it came from bare land, low vegetation density, and high vegetation density. An accuracy test for the interpretation of land use is performed through a quantitative calculation of the suitability of the field data. Based on the 100 random samples used, the accuracy of interpretation of the Landsat 5 TM image recording in 2009 amounted to 77 %. The time difference between the image recording field activities affects the value of the accuracy of the interpretation of land use. The four years between the image recording time and the fieldwork indicates changes in land use at the study site. Considering that most of the study sites were in urban areas, the city of Yogyakarta, there have been a lot of changes in land use in the period of 4 years, just as in the sample with coordinates 49M 435 285 9138025 which was originally bare land in 2009, but by the time of conducting survey activities had changed into grassland.

3.3 Comparison between Land Use and Temperature

Land use in the form of built-up areas had a higher temperature than another uses, in both 1992 and 2009. A temperature comparison between built-up areas and another land uses also showed a relatively large difference (Table 20.2). This is because the energy emitted by the built-up area tends to be greater than in other areas, as recorded by satellite sensors. Different uses of the roofs in the built-up area will affect the amount of radiance energy that is reflected from energy sources. A change in land use to built-up and urban areas will increase the temperature around and the impact on the temperature trend. Otherwise, land use in the form of high vegetation density (V2), has a low temperature compared with other land uses, both in the 1992 or in 2009. The increasing amount of V2 in urban areas would make the surrounding air became cooler, due to the evapotranspiration of the vegetation.

Table 20.2 Temperature difference between (1) built-up area (BA) and road (R), (2) built-up area (BA) and bare land (BL) (3) built-up area (BA) and low vegetation density (V1), (4) built-up area (BA) and high vegetation density (V2).

Years	Type			
	BA-R (1)	BA-BL(2)	BA-V1 (3)	BA-V2 (4)
1992	0.331429	-0.13751	2.089552	1.466195
2009	0.720502	0.361836	1.62579	2.081667

3.4 Correlation between NDVI and Temperature

A correlation test was devised to determine the relationship between temperature and the factors that affect its increase, and to determine the dominant factor in the land use in the temperature increasing process. The coefficient of determination with the symbol r^2 was the proportion of variability in the data that was calculated based on statistical models. The next definition states that r^2 was the variability ratio of the values that made the model with the variability of the original data values. r^2 was generally used as information regarding the suitability of a model. In the regression r^2 was used as a measurement of how well a regression line close to the value of the original data made models. If r^2 equalled 1, the figure showed that the regression line perfectly fit the data. Another interpretation was that r^2 was defined as the proportion of response variation explained by the regressor (independent variables / X) in the model. Thus, if $r^2 = 1$, it means that the model would be appropriate to explain all the variability in the Y variable. The first regression was performed in relation to the vegetation index and temperature. The results of the regression can be seen in Figure 20.3 and the results of the transformation of NDVI shown in Figure 20.4.

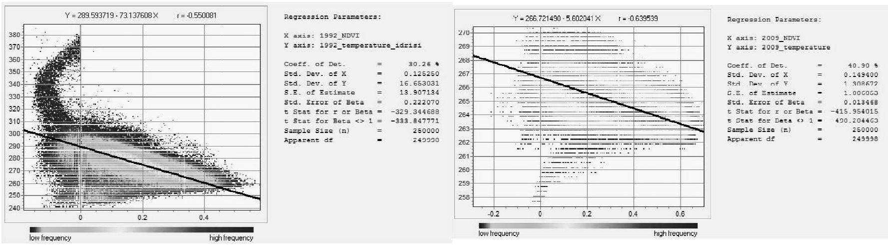


Fig. 20.3 Graphs of regression test of temperature and vegetation density in the (a) 1992 and (b) 2009

Figure 20.3 showed the coefficients of determination were 30.26% and 40.93%. The coefficient of determination (R) was negative, which means that both parameters are inversely proportional to both mean parameters, where the higher vegetation density, the lower the temperature. However, the value of R is small (below 0.5) which showed low relationships between temperature and vegetation density, thus suggesting that the value of vegetation density had less impact on the decrease in temperature. In addition, the value of the regression coefficient in 1992 became very small, at 0.3. This is due to pixels with NDVI values and high temperatures, and so different from the initial assumption that high NDVI causes a low temperature. This result may have been affected by a type of vegetation cover with less dense leaves so that heat absorption by oxygen was not optimal, and high vegetation density would still have high temperatures. This was not commonly found in 2009 due to land use changes that had occurred, which led to low NDVI values followed by a low surface temperature.

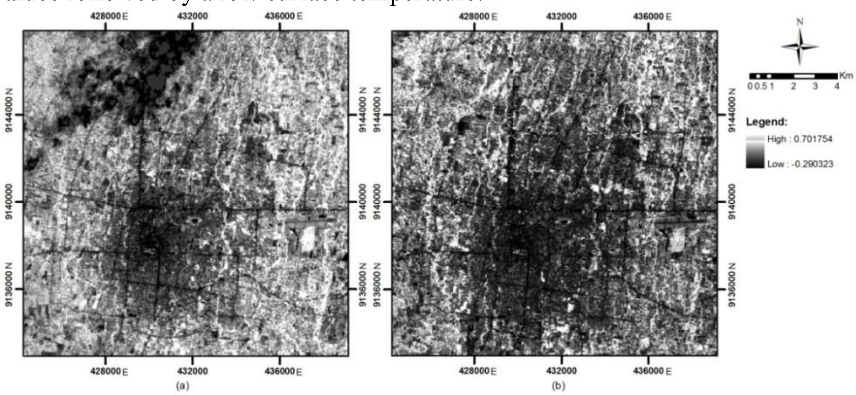


Fig. 20.4 NDVI distribution in (a) 1992 and (b) 2009

The relationship between the vegetation index value and the vegetation density in the field was obtained from the field measurement. The regression test (Figure 20.5) resulted in the the coefficient of determination (r^2) of 0.75.

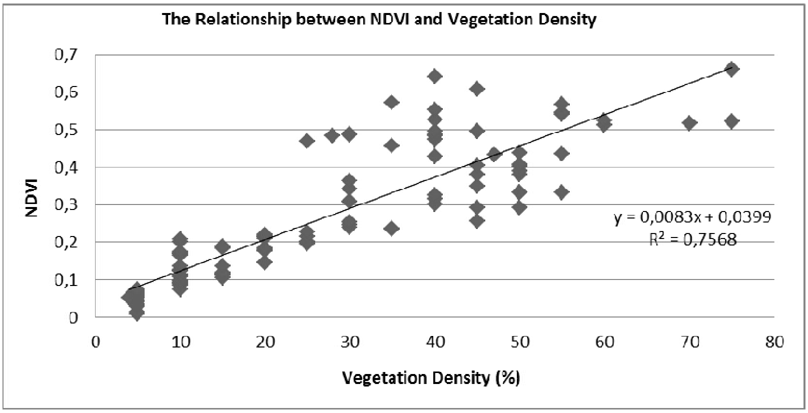


Fig. 20.5 The result of the regression test between NDVI and vegetation density in the field

A value of r^2 = 0.75 means that there is a relationship between the regression of (X) and Y variables. In this case, for example, if r^2 = 0.75 it means that 75% of the variation in the Y variable (dependent variable/response) can be applied to the X variable (the independent variable/explanation) while the rest is 0.25, influenced by unknown variables or inherent variability. Variability is the distribution of a set of specific values. Generally, the effect of variable X to Y is 75% while the other 25% is influenced by another factor. More data and further details were required to determine the factors that influence X and Y variables. A relationship test between variables was also performed for the whole pixel, and the results are given in Figure 20.6.

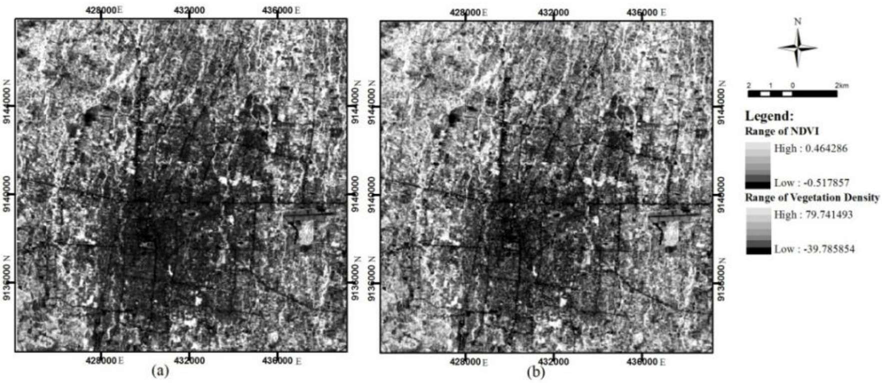


Fig. 20.6 Map of (a) NDVI in 2009 and (b) vegetation density in 2009 after application of y = 0.0083x + 0.0399 equation.

3.5 Correlation between NDBI and Temperature

In addition to NDVI, we also studied the effect of temperature fluctuations on the NDBI values in the study area**.** The relationship between the temperature and density of the building, based on the graph shown in Figure 20.7, has a fairly high correlation, and regression coefficients of 0.47 in 1992 and 0.44 in 2009. The values of the regression coefficient are positive, which means the temperature and density of the building are directly proportional: the higher the temperature the higher the density of the building. This can be affected by the type of roof that absorbs sunlight as an energy source for the re-emitted and be recorded by the sensors so the temperature rises on the image processing results.

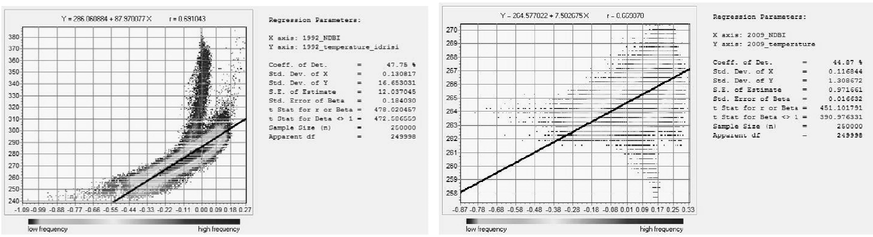


Fig. 20.7 Graph of correlation test between temperature and density of the building in (a) 1992 (b) 2009

A regression test was also carried out for NDBI and density of the building. The statistical model showed that the coefficient of determination is 47.76% so there was 50% influence on the rise of density in the building, with temperature. The type of building also affects the increase of surface temperature. Thus, it is not merely the density of the building that is a major factor in the rise of temperature. Details of the changes of density in the building can be seen in Figure 20.8.

After the sampling process and field measurement to see density of the buildings per pixel, the resultant graph of the relationships between NDBI and density of the buildings is as follows:

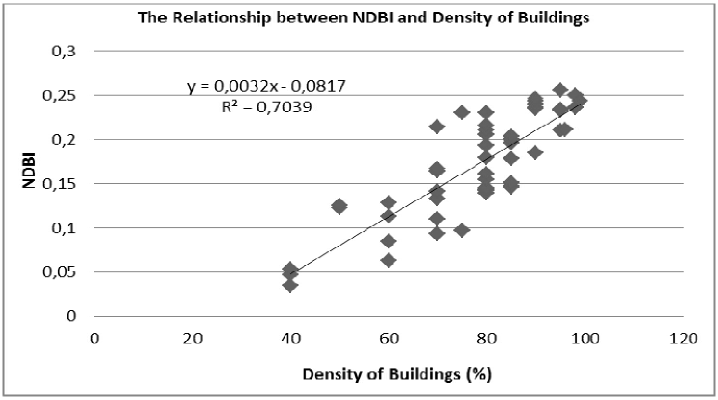


Fig. 20.8 The relationship between NDBI and density of buildings based on field measurement

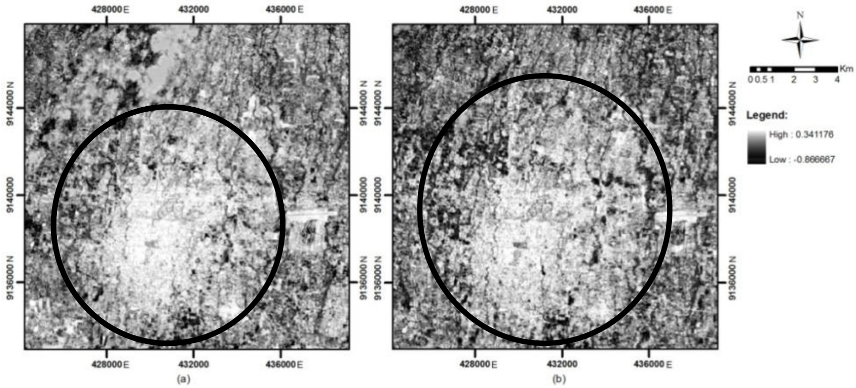


Fig. 20.9 Distribution of NDBI in 1992 and 2009

Figure 20.9 shows the broad changes in the density of buildings. As there is expansion in the type of land use settlement, so temperature rise is affected.

The relationships between NDVI and vegetation density, and also NDBI and building density, according to field measurement, were explained by the coefficient determination value. This coefficient value also shows that the pixel value (digital number/ DN) from Landsat TM can be used to examine the vegetation density or building density in the field. NDVI represents vegetation density and NDBI represents building density that can be compared with the temperature from Landsat TM.

4 Conclusions

This chapter presents the results of the temperature trend and the correlation with the land use/land cover in the Yogyakarta city. Processing shows that the distribution and trend of temperatures from the Landsat 5 TM thermal band for a period of 1992 and 2009 have risen an average of 0.99˚C as an indicator of the temperature trend in Yogyakarta city. The results of the relationship between brightness temperature and the land use/cover pattern (LUCP) is seen in the results of the transformation of NDVI and NDBI correlated with brightness temperature values. Correlation values of temperature and those two parameters showed that the value of NDBI transformation is the most dominant factor affecting the increase in temperature with a regression coefficient reaching 0.5184.

Land use in Yogyakarta city is classified as high vegetation density, low vegetation density, built-up land, and bare land. The spatial pattern of land use in 1992 was dominated by vegetation while in 2009 was dominated by built up areas. The relationship between the temperature trend and land use change patterns seen by quantitative methods is based on the temperature rise of every land use object.

References

Betchel B, Zaksek K, Hoshyaripour G (2012) Downscaling land surface temperature in an urban area: a case study for Hamburg, Germany. Journal of Remote Sensing 4: 3184-3200

BPS (2012) Daerah Dalam Angka Kota Yogyakarta Tahun 2012. BPS, Yogyakarta

Chander G, Markham B (2003) Revised Landsat-5 TM Radiometric Calibration Procedures and Postcalibration Dynamic Ranges. IEEE Transactions on Geoscience and Remote Sensing 41: 2674-2677

Chen Y, Wang J, Li X (2002) A study on urban thermal field in summer based on satellite remote sensing. Journal of Remote Sensing for Land and Resources 4: 55–59

Chen, Xhiao-Ling, Zhao HM, Li PX, Yin ZY (2005) Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sensing of Environment 104: 133-146

Cox JR, Hodges R, Rozenweig C, Solecki WD (2005) Skin of the Big Apple: Characterizing the Surface Heat Island of New York City and Integration with MM5 Climate Model. Paper presented at EPA Conference Call, 26 January 2005

Daldjoeni N (1986) Pokok-pokok Klimatologi. Penerbit Alumni, Bandung

Danoedoro P (1996) Pengolahan Citra Digital. Faculty of Geography Universitas Gadjah Mada, Yogyakarta

Danoedoro P (2012) Pengantar Penginderaan Jauh Digital. Andi Press, Yogyakarta

Deosthali V (2000) Impact of rapid urban growth on heat and moisture islands in Pune City, India. Atmospheric Environment 34: 2745-2754

Friedl M A, Broadley CE (1997) Decision tree classification of land cover from remotely sensed data. Remote sensing of environment 61: 399-409

Gallo K P, Owen TW (1998) Assessment of urban heat island: a multi sensor perspective for the Dallas Ft Worth USA region. Geocarto International 13: 35-41

Hawkins TW, Brazel AJ, Stefanov WL, Bigler W, Saffell EM (2004) The role of rural variability in urban heat island determination for Phoenix, Arizona. Journal of Applied Meteorology 43: 476-486

Kopec RJ (1970) Further Observations of The Urban Heat Island in A Small City. Bulletin of the American Meteorological Society 51: 602–606

Lillesand TM, Kiefer RW (1997) Remote Sensing and Image Interpretation 4th Editions. John Wiley and Son, New York

Lin CY, Chen WC, Chang PL, Sheng YF (2011) Impact of the urban heat island effect on precipitation over a complex geographic environment in Northern Taiwan. Journal of Application Meteorology Climatology 50: 339-353

Liu L, Zhang Y (2011) Urban heat island analysis using the Landsat TM Data and ASTER data: A case study in Hong Kong. Journal of Remote Sensing 2011 3: 1535-1552

Owen TW, Carlson TN, Gillies RR (1998) An assessment of satellite remotely-sensed land cover parameters in quantitatively describing the climatic effect of urbanization. International Journal of Remote Sensing 19: 1663-1681

Purevdorj TS, Tateishi R, Ishiyama T, Honda Y (1998) Relationship between percent vegetation cover and vegetation indices. International Journal of Remote Sensing 19(18): 3519-3535

Rao PK (1972) Remote sensing of urban "heat islands" from an environmental satellite. Bulletin of the American Meteorological Society 53: 647-648

Saaroni H, Eyal BD, Arieh B, Oded P (2000) Spatial distribution and microscale characteristics of the urban heat island in Tel-Aviv. Landscape and Urban Planning 48: 1-18

Streutker DR (2002) A remote sensing study of the urban heat island of Houston, Texas. International Journal of Remote Sensing 23(13): 2595-2608

Streutker DR (2003) Satellite-measured growth of the urban heat island of Houston, Texas. Remote Sensing of Environment 85: 282-289

Sutanto (1994a) Penginderaan Jauh Jilid I. Gadjah Mada University Press, Yogyakarta

Sutanto (1994b) Penginderaan Jauh Jilid II. Gadjah Mada University Press, Yogyakarta

Tian Q, Min X (1998) Advances in study on Vegetation indices. Advance in Earth Sciences 13(4): 327-333

Weng Q (2001) A remote sensing-GIS evaluation of urban expansion and its impact on surface temperature in Zhujiang Delta, China. International Journal of Remote Sensing 22(10): 1999-2014

Wang J, Rich PM, Price KP, Kettle WD (2004) Relations between NDVI and tree productivity in the central Great Plains. International Journal of Remote Sensing 25(16): 3127-3138

Zha Y, Gao J, Ni S (2003) Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing 24(3): 583-594

Zhao-ming Z, Guo-jin H, Rong-bo X, Wei W (2006) Land Surface Temperature Retrieval of Beijing City using MODIS and TM Data. Proceeding of Geoscience and Remote Sensing Symposium, 2006. IEEE International Conference. Beijing, China, July 31 2006-Aug. 4 2006: 1094 – 1096

Environmental Data Visualisation EnVIS – Linking Real-Time Sensor Data to Spatial Data Infrastructures for Web-Based Visualisation

Christophe Lienert[1] and Stefan Meier

Abstract This paper illustrates the concept and first achievements of the multi-annual Environmental Data Visualisation EnVIS project. The project is a response to the growing expectations of society for access to data from sensor networks via browser-based and mobile device-based maps. It deals with joint presentations and publications of various inner-cantonal time-series data originating from operational environmental measurement networks (hydrology, meteorology, agronomy) on a public web-based visualisation platform. Technically, it deals with the integration of real-time measurement data from environmental sensor observation networks into administrative spatial data infrastructures (SDI) and geographical information systems (GIS) for the purpose of real-time cartographic visualisation on the web. Internal administrative collaboration between various environmental specialist units and GIS-specialists is driven by the need to visualise real-time time series data online. . The focus in this paper is on the project structure, envisaged system architecture, the data model, interoperability issues, and different aspects of the web-based cartographical user interface.

1 Introduction

EnVIS is an acronym for environmental data visualisation. It represents a multi-annual project mandate to conjointly manage and visualise various environmental data that is being collected in (near) real-time in various departments and divisions within the administration of the Swiss Canton of Aargau, the fourth most populated state in Switzerland. The political mandate for EnVIS was conferred by the Governing Council via the Secretary General of the canton's Department of Construction, Traffic and Environment (in German: Bau, Verkehr und Umwelt

C.Lienert
Canton of Aargau, Department Construction, Traffic and Environment, Division Landscape and Waters, CH-5001
Aarau, Switzerland
e-mail: christophe.lienert@ag.ch

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_21,

(BVU)) in June 2013. Financial means for EnVIS have been approved for a period of two and a half years and earmarked within specific GIS/IT-budgets.

Overall project management resides with the canton's hydrometry unit, a unit that traditionally deals with questions related to measuring components of the hydrological cycle. This task includes various generic questions, such as those involving the collection, transmission and central management of real-time environmental data, including precipitation, river discharge and groundwater. The unit manages a network of more than forty full-fledged, fully automated and developed river discharge gauging stations, four semi-automated flood retention basins, and a series of online rainfall gauges. The unit's main mandate is to monitor and forecast superficial hydrological elements, in order to provide internal and external customers with high quality time series data for different purposes, such as flood/drought management, water ecology, or hydropower utilisation.

Currently, several organisational units within the BVU department are collecting environmental data (e.g., soil moisture, glazed frost warnings, potable water, air quality), but do so in a rather fragmented manner, due to different data ownership, different measurement parameters, and different subject-specific requirements. The first internal surveys have showed that a closer linking of data and an enhanced interface management for more integrative environmental data processing workflows would be of mutual benefit.

1.1 Project Structure

Various data owners and data sources, but also technical and organisational traditions and mentalities had to be considered and addressed when the first project presentation meetings were held. All sections and departments involved were invited and had to be convinced of the EnVIS concept, and about the advantages of sharing and disseminating corporate data for visualisation in a joint web-based portal. The project structure in Figure 21.1 shows the head of the BVU department as the contracting authority. Two division heads constitute the steering committee and one section head is the responsible project manager with additional steering tasks. Overall project management resides with two unit heads who have the necessary technical and organisational expertise and who manage public procurement with external contractors, the deliverables, quality and costs.

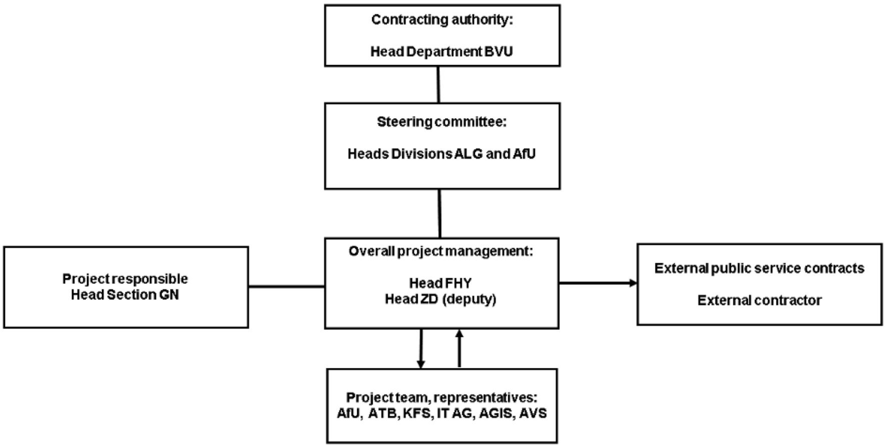


Fig. 21.1 Project structure of EnVIS (BVU = Department of Construction, Traffic and Environment, ALG = Division of Landscape and Waters, AfU = Environment Division, GN = Section Water Resource Management, FHY = Hydrometry Unit, ZD = Central Services Unit, ATB = Division of Civil Engineering, KFS = Cantonal operational headquarters, IT AG = IT Services Canton Aargau, AGIS = GIS Service Centre Canton Aargau, AVS = Division of Consumer Protection).

Although EnVIS was primarily intended to become a visualisation platform enabling the public to access real-time data, internal users are of course also addressed. Among the internal target group are members of the cantonal operational headquarters (KFS, see Figure 21.1) who are highly dependent on real-time measurement data as they hold the managerial responsibility for natural hazard, emergency and crisis response, and intervention planning.

1.2 Questions, Goals and General Framework for EnVIS

Goals for the projects were formulated during project submission. Further specification of the goals (mentioned below) and comprehensive variance analysis are to be worked out by an external contractor. Since cantonal GIS/IT projects are commonly contracted in a two-tiered manner (first feasibility, then main project), the feasibility study tackles questions about how to handle large amounts of real-time measurement data, and how to link and process these with spatial information stored in existing spatial data infrastructures (SDI) in order render them as interactive, real-time GIS-data (Choi et al. 2005, Kubicek et al. 2013, Lienert et al. 2011). Particularly important are geo-processing workflows that interpolate intermittent point data to surfaces (Iosifescu et al. 2010). Fundamental technological questions must also be addressed concerning the kind of geo-processing and visualisation software that is best suited to concurrently deliver such data, and how to embed existing, or even newly acquired, technology in today's cantonal IT system topography. Further goals include:

- Integration of environmental data for a holistic view of different complexes of themes (e.g., flood management, environmental monitoring)
- Establishment of an organisational structure for rapid integration of spatio-temporal data from different internal data-owners within one system
- Optimisation of data flows, ranging from capture and analysis to web-based publication of existing environmental data
- Creation of a multi-media visualisation interface, based on data originating from various proprietary database management systems.
- Delivery of service-oriented cartographic visualisation products for public/society (primary user) and for internal users (secondary users)

Existing infrastructures within the cantonal administration, day-to-day operations and future related projects require a general framework for EnVIS, consisting of the following items:

- Use of all existing and future time-series data that have "spatio-temporal relevance" (i.e., real-time as well as near real-time of up to 1-3 days)
- Use of existing, centralised SDI infrastructure (close cooperation with the Cantonal GIS Service Centre AGIS)
- Application of new and pioneering technology and concepts (cloud-based data storage, web services for retrieval, processing and visualisation)
- Evaluation and interface management for planned cantonal strategy eGovernment, with tailored citizen-oriented services.
- Evaluation and interface management to planned cantonal statutorily regulated implementation of geo-information law

The linkage of real-time data to the SDI infrastructure is a new requirement for the Cantonal GIS Service Centre AGIS. A board, consisting of GIS specialists in each division (AGIS Board), periodically assigns priorities for future works and developments. When the EnVIS project began, the requirement of "real-time in GIS" could be triggered and is now high on the board's agenda. Web GIS and mobile GIS are additional new GIS requirements put forward to the Cantonal GIS Service Centre by the EnVIS project. More detailed specifications for these requirements are given in the next section.

2 GIS-Data Model for Real-Time Data, Architecture and Cartographic User Interface

The core technical problem of the EnVIS idea is the development of a generic, sustainable data model for real-time data and its application in GIS-environments (Hoel 2008, Richardson 2013, Wang et al. 2013). Geodata is “spatial data having a specific temporal reference with which the extent and properties of specific spaces and objects are described, particularly their position, state, use and legal relationships” (GEOIG 2010). An observation, in contrast, is an “action whose result is an estimate of the value of some property of the feature-of-interest, at a specific point in time, obtained using a specified procedure” (INSPIRE 2011, Rautenbach et al. 2013). Basically, real-time data management has to be built on the following basic entities:

- Registry
- Objects
- Spatial representation
- Cases
- Dates
- Measurements
- Methodology

The concept of the overall system architecture may be viewed as a four-tiered system. As shown in Figure 21.2, the basis is set by the data owner. The next elements are the joint data storage, analysis tools, GIS linkage and visualisation.

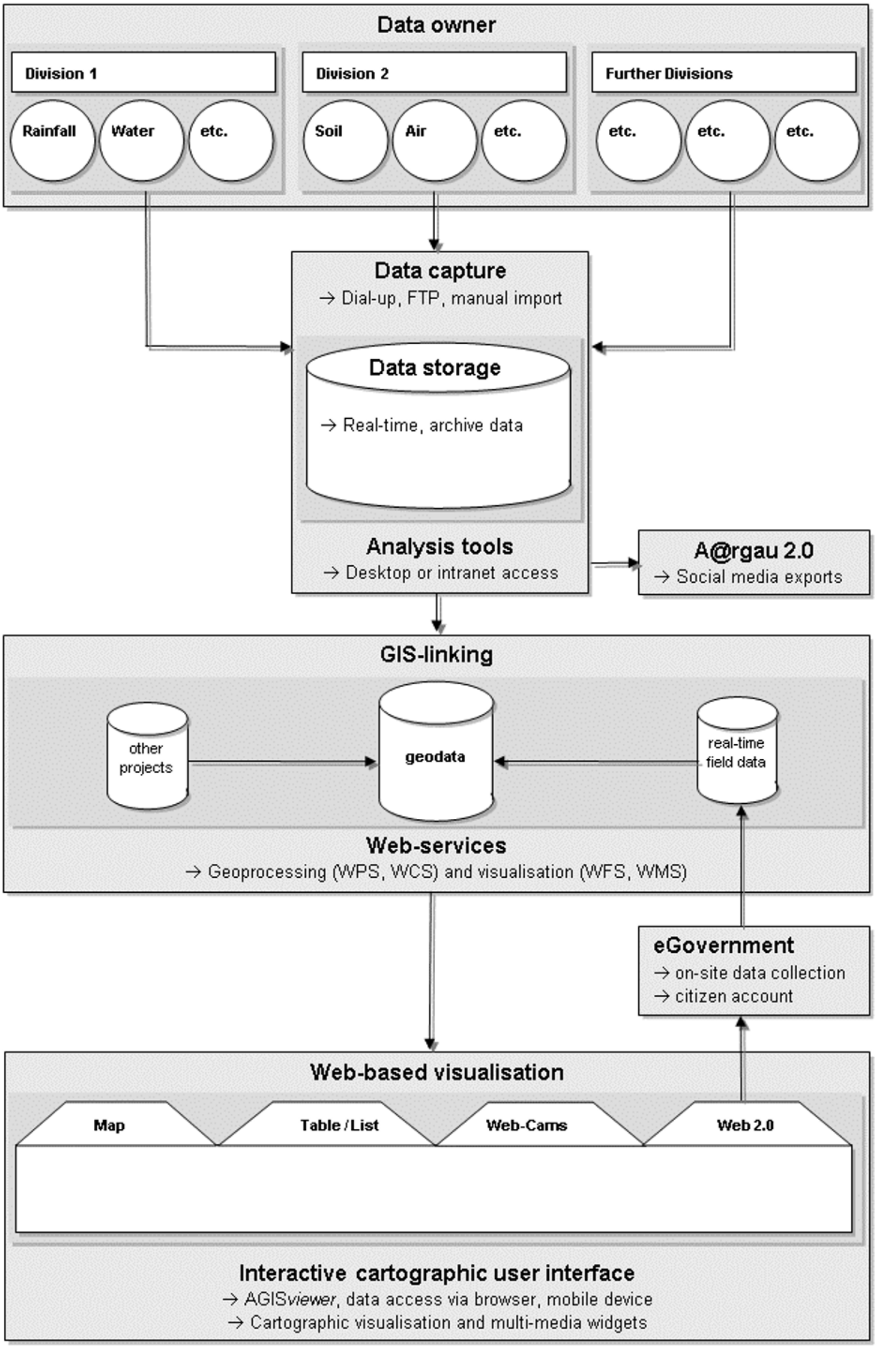


Fig. 21.2 Concept of the overall EnVIS system architecture

2.1 Data Owners

Data owners contribute to EnVIS with their raw or partly processed data. This data is either directly integrated, or integrated via the import/export interfaces of the data owner's proprietary system. Each data owner is free to use their own data management systems for storage, analysis and reporting. EnVIS is simply an offer for data owners to visualise data conjointly with other data owners. To do so, master data storage is established, which is either fed by automatic data imports from the data exports of the owners' preferred database management system, or fed by data from direct gauge retrieval.

2.2 Data Capture

Data capture requires that data is digitally available in real-time and accessible over dial-up networking on the respective data logger, or on FTP-servers in or outside the Aargau canton. Data must be gathered or delivered in a common syntax and format.

2.3 Data Storage

Data storage takes place in a dedicated commercial time-series database management system which is directly linked with two further modules: a data retrieval terminal and a distributed service manager that simplifies communication and data transfer between the relevant processes within the time-series database. Real-time and archive data are stored in this database, while data access takes place via desktop-based clients or intranet portals.

2.4 Analysis tools

Various analysis tools are an integral part of the commercial time-series database management system. Automated data validation and correction, as well as various sorts of data analysis, are provided, along with reporting frameworks. Depending on the needs expressed by the EnVIS project partners, further reporting tools may be developed and ordered from the software vendor.

2.5 Interface A@rgau 2.0

Before data is provided with a spatial reference (i.e., by linking it with GIS data), they may – for example if they meet a certain condition such as reaching a threshold – be dispatched to the public on social media channels. Since 2011, the Aargau canton has pursued a social media strategy called A@argau 2.0, in which new means of communication with the public are examined. In terms of EnVIS, a registered social media user may receive specific text messages related to certain real-time measurement values. Twitter messages are suitable to this end. Twitter is already being used in combination with operational measurement networks and crisis management (Terpstra et al. 2011, Yates and Paquette 2011).

2.6 GIS linking

Linking real-time time-series data constitutes a new and innovative requirement for the Cantonal GIS Service Centre AGIS, a project partner of EnVIS. It is intended to directly link dynamic measurement data with underlying geometry data. New spatio-temporal datasets with additional attributes will be generated. This entails a new service-oriented architecture (SOA) being set up, or existing ones being improved, so that GIS data models may reproduce the ever-changing spatio-temporal state and may be eventually visualised. The linking of time-space-attribute data should take place within the existing SDI infrastructure. By means of web-services (e.g., processing, feature, map services) and dedicated tools, functional requirements are formulated in terms of data workflows. These will be thoroughly tested in offline environments and later put online. These newly developed or refined web services are responsible for the automated, real-time generation of web-enabled, interactive EnVIS visualisation.

2.7 Web-based Visualisation

Access to the visual output generated from the above GIS linking is granted to the public or internal users by means of a (carto-) graphic user interface on the web browser or on mobile devices. Visualisations are characterised by plainness, cartographic compliance, interactivity, and by additional information windows, so-called widgets, containing lists, tables, graphs and multi-media content. The visualisation component of EnVIS will most likely be realised within the existing AGIS*viewer* infrastructure. This viewer based on ESRI-GIS technology is a product of the AGIS Service Centre and established within the entire cantonal administration. Many technical, functional and organisational achievements have already been accomplished and thus it is appropriate for EnVIS to be further built on this viewer.

2.8 Interface E-Government

At a later stage, the above mentioned AGIS*viewer* is also supposed to become the portal for users who actively collect field data and make it available in real time to other users, or for instantaneous visualisation for themselves. Applying this bi-directional concept (i.e., EnVIS as a means to consume spatio-temporal content, but also to create such content) addresses such Web 2.0 ideas and manifold possibilities open up related to GPS-enabled mobile devices for the collection of location-based, real-time field data, or more generally "citizens as sensors" (Engler et al. 2014, Foerster et al. 2011, Goodchild and Glennon 2010, Konecny 2012). These potentials are considered in the feasibility and may be partly implemented in the EnVIS project.

Collaborative data collection to obtain a more dense data basis and the involvement of citizens in this process, promotes the interface to the current cantonal E-Government project. With the planned "citizen account" authentication mechanism, which is one of the deliverables of the E-Government project, data quality and integrity could be ensured for such an envisaged extended Web 2.0 EnVIS visualisation platform.

3 Web Integration

Real-time data from data storage (see Figure 21.2) is combined with the AGIS-*viewer* using a web interoperability which was delivered with the commercial time-series database management. Additional data workflows within the AGIS-*viewer* result in a map, shown in Figure 21.3. Pre-built background maps can be selected from a large AGIS*viewer* catalogue. A specific map module "Hydrometry" is being developed and for the moment, linkage between time-series data and GIS data is confined to (gauge) points. Points in the map are coloured and symbolised so as to represent and differentiate various parameters (river discharge, precipitation), data owners and multi-parameter gauges (e.g., measurement of water level and water temperature at the same site).

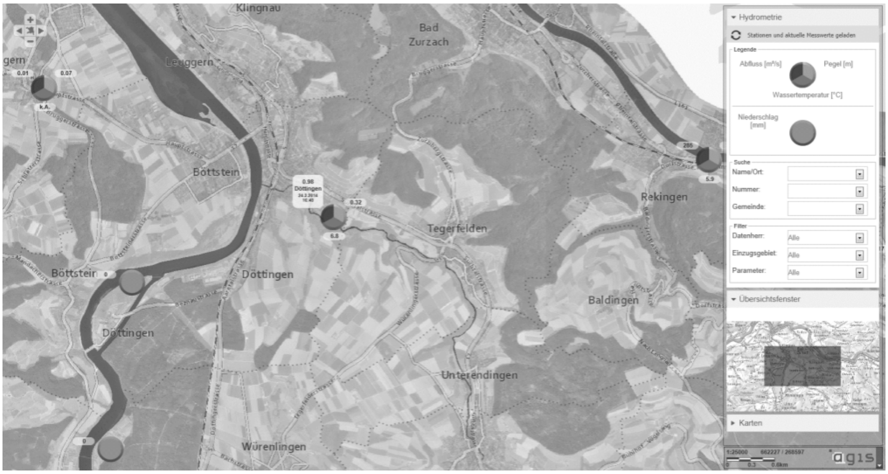


Fig. 21.3 Screenshot of the "Hydrometry" map module within the AGIS*viewer*

Traditional map navigation functionality in terms of zooming, panning, searching and filtering is provided, as well as a 3-level zoom concept (global, regional, local) which is linked to some generalisation (number of visible point symbols) and functionality changes (clicks and mouse-overs).

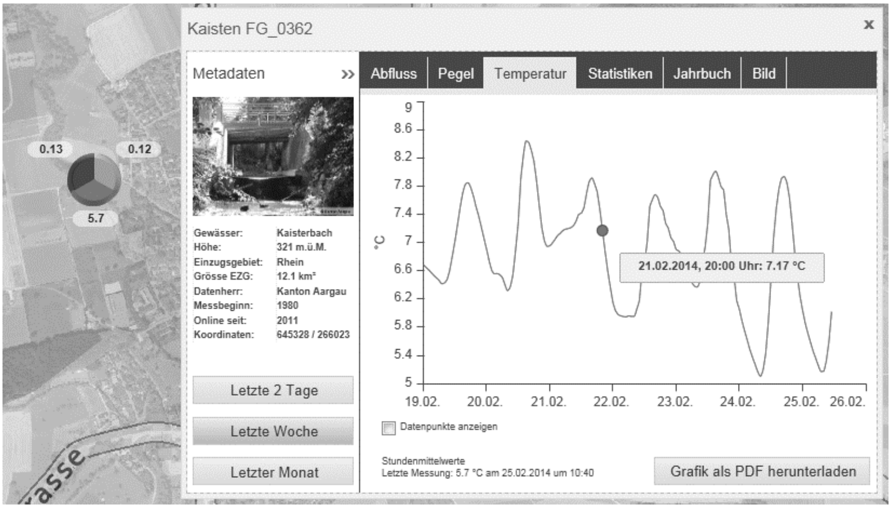


Fig. 21.4 Screenshot of the multi-tab window within the AGIS*viewer*

An additional window, shown in Figure 21.4, is based on several tabs in which time series graphs, photos, and the metadata of statistics are displayed. The window is movable and resizable. The time series graph may be further explored using a computer mouse.

4 Conclusions

EnVIS is a cross-departmental, inter-divisional project of the Aargau canton, Switzerland. It deals with the new and innovative management of environmental data, and focusses on their visualisation in a versatile, cartographic user interface. The primary user of EnVIS is the public, and its outcome is expected to be perceived as valuable by society, in allowing them to view environmental data in an integral, holistic way. Secondary users are internal specialists who need interrelated data for event-based analysis such as flood or drought management.

EnVIS – the amalgamation and joint visualisation of environmental data over different organisational units – clearly involves a considerable integration effort, that is functional and technical as well as organisational. However, the benefits for the project participant, and the added values (joint visualisation system, joint data management, joint basis for further data analysis and publication activities, access to several interrelated data parameters, event-based analysis, long-term integral monitoring, possible reduction of data management and database systems) justify these efforts.

With the linking of GIS and environmental measurement data, innovative new cantonal data products can be generated and existing data will be considerably enriched. The new requirements of the EnVIS project will stimulate activities in the field of real-time GIS, interactive web- and mobile-based data visualisation, and foster activities in the field of E-Government.

Last but not least, EnVIS will hopefully also help raise awareness about questions related to the environment and help sensitise society to the value of environmental sensors and data.

References

Choi JY, Engel BA, Farnsworth RL (2005) Web-based GIS and spatial decision support system for watershed management. J. Hydroinf 7(3):165–174

Engler NJ, Scassa T, Fraser Taylor DR (2014) Cybercartography and Volunteered Geographic Information, In: Fraser Taylor DR (ed) Modern Cartography Series, vol 5. London, Academic Press, pp 43–57

INSPIRE (2011) D2.9 Guidelines for the Use of Observations & Measurements and Sensor Web Enablement-Related Standards in INSPIRE Annex II and III Data Specification Development.

http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.9_O&M_Guidelines_V1.0.pdf. Accessed21 Jan 2014

GEOIG (2010) Swiss Federal Law on Geoinformation 510.62, GeoIG. http://www.admin.ch/opc/de/classified-compilation/20050726/index.html. Accessed 21 Jan 2014

Foerster T, Nüst D, Bröring A, Jirka S (2011) Discovering the Sensor Web through Mobile Applications. In: Gartner G, Ortag F (eds) Advances in Location-Based Services. 8th Intl. Symposium on Location-Based Services, Vienna, November 2011. Lecture Notes in Geoinformation and Cartography, Heidelberg, Springer, p 348

Goodchild MF, Glennon JA (2010) Crowd sourcing geographic information for disaster response: a research frontier. Int. J. Dig Earth 3(3): 231–241

Hoel E (2008) Data models in commercial GIS systems. In: ShekharS, Xiong H (eds) Encyclopedia of GIS, Springer, Heidelberg, pp 215–219.

Iosifescu-Enescu I, Hugentobler M, Hurni L (2010) Web cartography with open standards – A solution to cartographic challenges of environmental management, Env Mod and Software, 25(9): 988–999

Konecny M (2012) Early Warning and Crisis Management Cartographic and Geographic Information Research Agenda. In: Bandrova T, Konecny M, Zhelezov G (eds) Proceedings of the 4th International Conference on Cartography and GIS, vol 2, Sofia, Bulgaria, 2012

Kubicek P, Kozel J, Stampach R, Lukas V (2013) Prototyping the visualization of geographic and sensor data for agriculture, Comp and Electr in Agric, 97: 83–91

Lienert C, Weingartner R, Hurni L (2011) An interactive, web-based, real-time hydrological map information system. Hydrol. Sci. J. 56(1): 1–16

Rautenbach V, Coetzee S, Iwaniak A (2013) Orchestrating OGC web services to produce thematic maps in a spatial information infrastructure, Comp Env and Urban Sys, (37): 107–120

Richardson DB (2013) Real-Time Space-Time Integration in GIScience and Geography. Ann. Assoc. of Am. Geographers 103(5):1062–1071

Terpstra T, de Vries A, Stronkman R,Paradies, GL (2011) Towards a realtime Twitter analysis during crises for operational crisis management. In: Proceedings of the 9th International ISCRAM Conference, Vancouver, Canada, April 2012.

Wang X, Huang H, Ban Y (2013) Water and rainfall monitoring and warning system based on Flex GIS technology. In: Proceedings of the 21st International Conference on Geoinformatics, Kaifeng, P.R. China, 20–22 June 2013

Yates D, Paquette S (2011) Emergency knowledge management and social media technologies: a case study of the 2010 Haitian earthquake. Int J InfMgmt31(1): 6–13

Cartographic Issues of Volunteer Mapping of Environmental Noise Using Mobile Phones

Petr Duda[1]

Abstract The voluntary collection of geographic information is becoming increasingly widespread, even in the area of environmental data. At present there are applications that allow the acquisition of environmental noise data using mobile phones. The accuracy of these phones is very high (up to ± 1 dB), which allows the collection of reliable data and compilation of quality maps. However, volunteer users often do not have the time or motivation to perform long term static measurements, therefore a method to handle noise data obtained during the normal movement of a user in the field has been developed.

Solved problems include: a) dependence of the nature of the data on user behaviour, b) errors in the data, depending mainly on positioning error and c) questions about the adequate visualisation of that data. Analysis is supplemented with a view of the impact of noise and volunteer noise mapping on both individuals and society.

1 Introduction

As a result of the development of information and communication technologies, particularly in miniaturisation, data collection using handheld mobile devices (smart phones) is increasingly promoted in the field of geoinformatics. These devices already have sufficient computing power and a GNSS (GPS) receiver. This connection forms the basis of location-based services (LBS), which are used, for example, to search for locally relevant services, in social networking, crisis management, navigation etc. It has also become the basis for in situ data collection, made through sensors connected to the mobile phone.

One of these data types involves acoustic pressure (noise). Almost every mobile phone has a built-in microphone for voice communication. This microphone can be successfully used for noise measurement. At the present time some applica-

[1] P. Duda (✉)

Masaryk University

Brno, Czech Republic

e-mail: 147169@mail.muni.cz

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_22,

tions exist (e.g. NoiseTube (D'Hondt 2013); NoiseDrone (Nüst 2013) or NoiseSpy (Kanjo 2010)) which allow the measurement and sharing of noise data from mobile phones, and perform basic processing.

It is thus possible to create maps of noise pollution, both indoor and outdoor, using volunteer contributions even in vast areas, at relatively low cost. The nature of the noise data, obtained from these measurements, is entirely dependent on the behaviour of the persons engaged in the measurement, however, and interpretation of that data relies on its processing.

This chapter analyses the noise mapping technology that uses mobile phones and defines the basic issues in noise processing and the visualisation of data obtained using this method. It is divided into seven parts. The first part briefly outlines the impact of excessive noise on humans and society, the second and third parts are devoted to the technological fundamentals, Part Four discusses the behaviour of noise data originators - mainly volunteers, Part Five deals with influence of positional errors on the collected data quality and the sixth part discusses problems of visualisation. The last section outlines the possible future impact of this technology on individuals and society.

2 Motivation: Environmental Impacts of Noise

It is very difficult to accurately quantify the impact of noise, as some individuals are more sensitive to noise than others. However, excessive noise provably affects human health in many ways, and not only by causing physical damage to the hearing organ. Excessive noise, especially in protected housing areas, but also in the workplace or during a journey, causes stress, which in turn affects many other organs. As well as the heart, this includes the hormonal and immune system, and placenta or human foetal development. Excessive noise also affects quality of sleep, and has negative effects on the acquisition of language and reading in children, mental health, social behaviour and general human performance. Night noise may contribute to the causes of obesity, sleep disorders, subsequent work related injuries and shorten life expectancy (World Health Organisation 2007).

These problems concern a large part of the population. Excessive noise leads to the loss of approximately 1.6 million years of healthy life in the population of Western Europe (about 340 million) each year. The greatest threat to the environment, air pollution, leads to the loss of approximately 4.5 million years, so noise is the second greatest contamination (Theakston 2011). One hundred million people in the European Union (EU) are exposed to noise levels above 65 dB (World Health Organisation 2013). Excessive noise is the cause of approximately 3% of all deaths caused by heart failure (Theakston 2011).

Given that people tend to avoid long term occupation in areas affected by noise pollution, there is also an influence on the real estate market. Property prices for areas affected by noise levels of 65 dB are 10-22% lower on average than the prices of similar properties in quiet locations (Ogneva-Himmelberger 2010). With every extra decibel of noise, lease property prices are reduced by approximately 0.6 % (Baranzini 2005).

Noise is thus becoming one of the major factors influencing the attractiveness of location and health and well-being of an area's inhabitants. For this reason efforts towards systematic mapping and control have been under way for more than 40 years, implemented by several institutions, from strategic noise maps, environmental impact case studies, to the measurement of excessive noise by hygiene stations.

3 Measuring Noise Using Sensor Networks

Noise measurements are now routinely performed in various situations, including the validation of computational models or resolving resident complaints about excessive noise. Standard measurements are usually carried out over a longer period during the day and only in a few selected locations. Location as well as time is chosen so that the results can be generalised. Acoustic situations are likely to vary from place to place quite considerably (see chap. 6), which limits the validity of the resulting data to the immediate surroundings of the measurement point. For this reason, noise sensor networks, which facilitate the collection of noise data from larger areas, are frequently created.

The most common use of sensor networks for monitoring noise levels includes sensor networks monitoring aircraft noise around major airports, e.g. ANOMS (Brüel and Kjaer 2013). This system also enables on-line visualisation of measured noise data and noise sources (e.g. aircraft) within the map. It consists of a network of measuring stations (static and mobile), appropriately spaced around an airport and connected to a central data store and several workstations. It is also used to evaluate complaints from the residents of surrounding communities to aircraft noise. The mapping of noise is possible because the noise source location is always known (by radio telemetry and/or GPS) and there is almost no noise shadow. This allows the reasonably reliable calculation of data for other locations.

For representative results in an urban environment a large number of measuring stations is necessary. City Hall in Madrid (Manvell 2004) performed experimental measurements which resulted in the creation of a noise map. Acoustic data was recorded using several mobile monitoring units (cars with mounted sensors) and displayed on a map. In 2002, on the basis of 4395 measurements points, a noise map for the whole of Madrid was created. The system has been gradually improved and today is known as SADMAM. Among others, sensor systems monitoring noise include the system at Dublin Trinity College, which deals with problems of delayed communication (McDonald 2008).

4 Measuring Noise Using Mobile Phones

A similar method of data collection, utilised by Trinity College, can be used for sensor networks formed by smart (or site connectable) mobile phones. Laboratory experiments have shown that the built-in sensors of mobile phones have a surprisingly high accuracy in the measurement of acoustic noise levels, which gall below ± 1 dB in the case of correct calibration, a 1-second measuring interval (mode "slow"), and for a range of about 35-110 dB (Weber 2013, Stevens 2012).

Comparative field measuring with a second class noise meter (according to IEC 61672-1 norm) shows that the accuracy of a phone calibrated to measure noise against this type of sound level meter reaches ± 1 dB, mainly due to a wide range of inaccuracies not explained in laboratory (e.g. sampling frequency, microphone directivity, influence of observer, sound level meter case, variations in ambient temperature, variations in ambient pressure, wind, etc.). However, the main disadvantage of these devices is lower sensor dynamics and the resulting slower response, and they are not therefore suitable for the measurement of impulse noise. They are sufficient for the majority of noise from traffic, as well as many kinds of neighbourhood noise, however.

The response of a sensor depends, not only on the sound pressure, but also on the sound frequency (tone). The sensitivity range of a mobile phone microphone is usually slightly smaller[2] than that of a certified sound meter (Stevens 2012). This is critical for the assessment of, for example, traffic noise, where, in some cases, the tonal component is dominant at the edge of the spectrum (e.g. low frequency of freight car transport, or, conversely, a high frequency of rail transport). The result may also be distorted by built-in phone software for adjusting voice frequency[3].

5 Volunteered User Behaviour

User behaviour affects the measured data quite significantly. While the behaviour of hired workers is subject to regulations, this might not be the case for voluntary users. In principle it is possible to distinguish three kinds of volunteers whose data may have somewhat different characteristics:

1. Users collecting data during their own activities (commuting to work, school, sports, such as cycling, etc.). In this case, the nature of measured data is per-

---

[2] With approximately appropriate inaccuracy ± 5 dB is usually registered in the range of about 125 to 12 500 Hz, however, by appropriate calibration techniques may be acceptable range extended to virtually all audible frequencies and accuracy of ± 1 dB across the audible spectrum achieved, although at the edges of the spectrum is, even after, detected data uncertainty somewhat higher.

[3] So that the resulting voice is more understandable for humans.

sonal; mobile phones thus act as noise dosimeters (equivalent to radiation dosimeters). In this way it is possible to capture a noise situation as it affects an individual human.[4]

2. Users collecting data exclusively or in their free time. This may include people who are doing their measuring statically (e.g. a user considers a suitable outdoor or indoor location and conducts an objective measurement according to professional guidance and standards to ensure the smallest possible impact of undesirable noise sources), people who are trying to measure the loudness of their devices or people who, for example, simply put a mobile phone on a window sill. Again, the noise data is burdened with some errors, but these are primarily in terms of uneducated operation and the imprecision of the sound level meter itself.
3. The third group of volunteers are those people who are somewhere between the two previous cases. These could be participants of gamified version of collecting application. Methods of data collection can be controlled by a narrative game to some extent as shown in NoiseBattle (Martí 2012).

The main conditions that determine the quality of the resulting data are therefore the degree to which the user is interested, and how much time they are willing to devote. Although data collected during a user's personal activities is much less accurate, this method allows the collection of more data. Such data must be radically processed, however: subtracting the noise of user steps, removing unwanted noise generated by the device, adjusting the position of the individual noise samples and spatially assigning them to the noise source (road, track, industrial facilities, cultural events, etc.).

Some of these steps can be relatively easily automated (e.g. removing unwanted noise to mobile phone by monitoring its activity, and the calculation of phone orientation errors using data from the GPS), other automation requires considerable effort (e.g. to determine the noise level of actual steps it is necessary to also record the audio track and then analyse it, for precise positioning at the level of ± 1 m is necessary to install a different positioning system, eventually performing a comparison of the saved route with a database containing information about the course and widths of roads, paths and traffic lanes).

It is possible to limit all the above errors, at least to some extent, using information that the user appends to measured data during or after measurement. But greater demands on a user's attention and time reduce the incentive to collect this additional data or any measurements at all (such a user is gradually moving from the first group of users to the second).

As shown in some studies (e.g. Deterding (2011)) gamification can significantly contribute to the use of application features and attract new users. The basic

[4] With some restrictions, such as that the noise level at the average height of the human ear is somewhat different from the noise at waist level, where people usually wear mobile phones during measurement.

concept of gamification in collecting noise data using mobile phones has been developed in the NoiseBattle project (Martí 2012).

The principle consists of conquering individual grid cells by measuring the noise based on the frequency and quality of measurement (game cells cannot be confused with cells to evaluate noise levels). This concept can be extended by adding the possibility of enhancing the quality of measurement by entering additional information (e.g. for the measurement of noise emitted by a user's own walk in a certain kind of shoes, about crossing the street during measurement, or manually entering the course of a route based on satellite imagery after the measurement).

6 Effect of Positional Inaccuracies in Outdoor Measurements on Noise Data Quality and Visualisation

The basic problem in using mobile phones as a fully-fledged sound meter outside a building is determining their position at the time of measurement. Nowadays, mobile phones are commonly equipped with a GPS receiver, which this task somewhat simplifies.

A mobile phone's GPS is frequently subject to high inaccuracies in positioning, especially in urban areas. The main cause is a reflected GPS radio signal, which leads to an incorrect calculation of transmitter-receiver distance and subsequently to incorrect positioning. The resulting inaccuracy in an urban environment is usually between 3-40m (Modsching 2006).

In some cases, therefore, the location of the resulting measured points is automatically adjusted (e.g. when measured on the street, but the location data indicates location in built-up areas). Such results, however, have limited relevance.

GPS receivers are also a significant consumer of power and can't be used everywhere, so other positioning methods have been developed. They are usually based on triangulation between the base stations of a cellular mobile network or Wi-Fi stations (Rekimoto 2006, Borrielo 2005). These methods are used, for example, for indoor navigation, and are also suitable in cases where the use of a GPS signal fails (e.g. in dense forest, street canyons, etc.) and where there is not such strong emphasis on positioning accuracy. This is not the case in noise measurement, which require the position from the source at least in the order of decimetres for calculations.

The measured noise level greatly depends on the distance from the noise source. In the case of a noise source with sound power *P*, emitting noise equally in all directions, at a distance *r* from the imaginary centre of the acoustics, emission is the medium intensity of sound given by:

$$I_{str} = \frac{P}{4\pi \cdot r^2} \quad (1)$$

If L_{p1} is sound pressure level at a distance r_1, and L_{p2} sound pressure level at a distance r_2 from a point source, it is possible to express the relationship between the two levels as:

$$L_{p2} = L_{p1} + 20 \cdot \log\frac{r1}{r2} \quad (2)$$

Thus, theoretically, for a point source, sound pressure level decreases with the square of distance ($1/r^2$). In the case of a linear source, such as heavy traffic on a road or a passing train of wagons, acoustic waves are in the shape of concentric cylinders; thus the sound pressure level theoretically decreases proportionally to the distance (1/r). In real environments, however, there is further attenuation, due, for example, to absorption in the air, fog, rain, snow, wind, temperature gradients, atmospheric turbulence, ground effect or obstacle influence.

The graph in Figure 22.1 shows that the usual position error of GPS receivers can cause significant problems when interpreting common noise conditions.

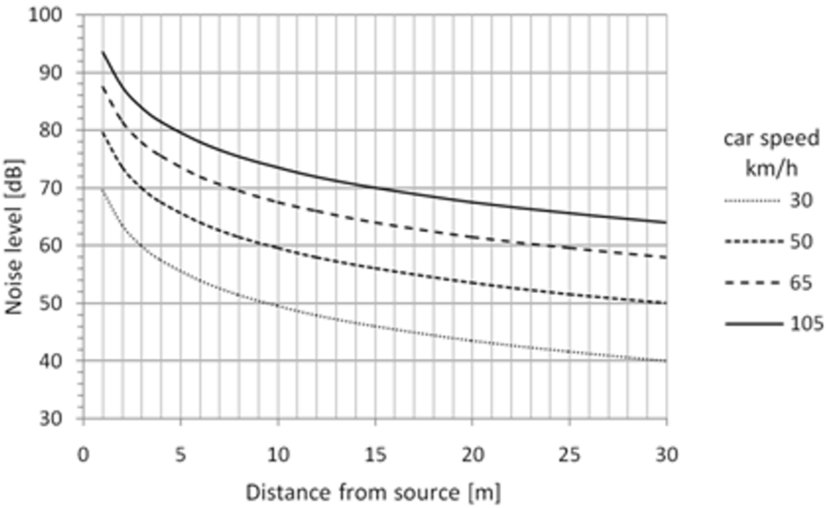


Fig. 22.1 Dependence of noise levels typical of urban noise sources (individual cars at designated speeds) decrease at a distance from the noise source

Applications for collecting noise data from mobile sensors therefore sometimes use the principle of identifying the measured point to the nearest point in a street network (to a certain limit of values, as in NoiseTube (Stevens 2012)). Applications do so automatically and do not offer the ability to tell the application whether a user is continuing down a street or has already moved away. So there may be significant errors in data interpretation from open spaces such as parks, gardens or squares.

Normal user movement along a street network also puts considerable demands on the accuracy and detail of the supporting geodatabase. Ordinary pedestrians are usually not moving in the middle of the road, but rather on pavements. In the case of an asymmetrical pavement layout or overly wide street there may be inaccuracies. When measuring noise from pedestrians, a database should also contain the lines of sidewalks, and if possible, an indication of their width, and also crossings and traffic lights.

The collecting application should then contain functionality through which a user could specify which side of street they are on and if he changes it. Final processing should then interpret each measurement point in the context of other points and assign whole point sequences to lines only, to avoid "jumps" from some points to adjacent lines.

For cities with block development it is appropriate to supplement the database with ground plans of buildings, because, when measuring outdoor, the measuring mobile device cannot be located in the building. It should be noted that an assignment of the measuring position on the line of a road or pavement causes an entirely new kind of inaccuracy, which is only indirectly dependent on the inaccuracy of distance measurement, but also on the quality of the data from a spatial database.

7 Visualisation of Outdoor Noise Data from Mobile Phones

Current outdoor environment noise maps can be divided into three kinds. Areal maps which indicate a) the average value of the noise level for each grid cell, and b) areas with noise levels between specified limits (noise contours). These types of maps indicate noise distribution in an area and are currently the most commonly used. Linear maps show the noise levels at certain sections of linear noise sources (e.g. road or rail). These maps usually give the noise level at a standard distance from the source. Point maps indicate the level of noise on measured or calculated sites. These three types of map can also be combined.

7.1 Determining the size of computational and visualisation grid cell

Existing applications exploiting noise measurements primarily use areal or point maps of type. The NoiseTube application (Stevens 2012) performs (after linearization to the street network) averaging of noise into grid cells at the edge of 20m. As is evident from Figure 22.2, there may in this case be some inaccuracy in the display and interpretation on contact of more streets. The data in one cell can commonly have a range of 10dB or more, depending only on the number of measurements on the paths perpendicular to the incoming linear noise source or the roads located behind the noise barrier.

A grid cell with an edge measuring 20m, although it enables relatively fast data processing and shows the generalised situation without further necessary modifications, can thus only be used in areas with a regular street network and block development, where there are fewer differences in the noise situation. The experiments performed also shows that the aggregation of cells of this size does not provide better information about the noise situation and creates a rather broken image, which can often not be used as a reliable basis for determining the location of the cause of a noise increase (or decrease).

When using smaller cells the situation no longer occurs. An acceptable cell size appears be 4 to 10m. Smaller sizes do not seem to bring a significant improvement in spatial accuracy. In a tested sample of 5600 measurements, obtained when walking in an area of low and middle height buildings in December 2013, the average deviation of points from a precise route about 3m[5] has been measured. If we add deviation from the actual position of the pavement, to an approximated line in the database, which occurs during normal measuring, the values may continue to increase.

Grid cell size can still be affected by the speed of the sensor or the average width of the street.

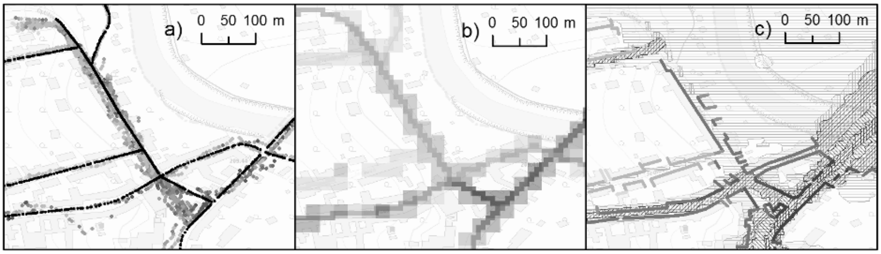


Fig. 22.2 Interpretation of the measurement of noise using mobile phones a) assignment of measurement points to lines, b) generalisation of values by grouping to individual cells and by application of zonal function, c) calculation and visualisation of line noise emissions (in the calculated background strategic noise map, which handles only the southern area) (Sources: Own work, ČÚZK, CENIA).

7.2 Visualisation of noise by lines

It has recently been quite common to represent linear sources of noise (road, rail) by means of linear maps. A significant advantage of this method is the ability to easily generalise places with the same noise conditions, and the resulting visualisation is more evident. In a similar manner is also possible to visualise line emission zones - pavements, pathways and facades. The size of line cells may be determined in a similar manner, as in the case of area maps.

[5] Average error 2.91m, median 2.44m, standard deviation 2.94m for log-normal distribution

Noise levels at house facades may be obtained in two ways: a) by measuring on the facade level (and correcting for noise reflection), b) by calculation with knowledge of the noise situation on the pavement and mutual distances of the pavement, noise source and the facade. Thus, in the case of noise from road traffic in urban streets, it can often be relatively easy to obtain information about the actual noise exposure of inhabitants.

7.3 Visualisation of measurement uncertainties

Noise data measured using mobile phones is affected by two basic types of uncertainty: a) uncertainty of immediate measurement, b) uncertainty about the temporal variability of the noise source.

Both uncertainties can be quantified in units of measurement (dB). The second uncertainty can be quantified in units of measurement only if two things are known in advance: a) possible range of representative values, and b) minimum measurement time for each noise source phase operating. For example, for general road transport, each specific phase of the traffic situation (e.g. morning peak) needs to be measured for at least an hour. When there is more traffic (more than 100 vehicles per hour) it is possible to shorten the representative time.

The visualisation of the combination of these two uncertainties may therefore occur if an approximate range of noise values is known. If not, it is possible to identify uncertainty as a proportion of the measurement time and the time required for representative measuring, and visualise it as a third variable.

Uncertainty can also be determined for traffic noise using the equation:

$$u_{sou} \cong \frac{C}{\sqrt{n}} \text{ [dB]}, \qquad (3)$$

where n is the number of passing vehicles and C the weighting coefficient (in the case of road transport with mixed traffic $C = 10$). In this case it is necessary to include data about the number of passing vehicles.

8 Conclusion

Environmental noise today represents a serious problem for civilisation. Measurement using mobile phones offers the potential to not only obtain sufficiently reliable alternative data sources, for example in areas that are not covered by strategic noise mapping (see European Union 2002), but also to raise public awareness about this dangerous, but ubiquitous phenomenon. The basic concept of measurement can be extended and refined in many ways. First of all it is necessary to resolve the question of the visualisation of uncertainty, because the noise data from mobile phones is heterogeneously affected by uncertainties, due to different user behaviour, position errors and errors in actual noise measurement.

Noise measurement using mobile phones cannot substitute for professional sound level meters and calculation applications. However, they are becoming a common platform for testing the whole concept of the voluntary collection and processing of environmental data. If this model expands and improves, it may be possible to significantly reduce the gap in communication about environmental noise between national, regional and local governments and citizens. In the near future it will not be necessary to deal with most cases of excessive noise long after the limits was exceeded. Using these methods it will also be possible to extend noise mapping to the countries which do not have enough resources to create these maps with the help of expensive professional technology.

References

Baranzini A, Ramirez J (2005) Paying for quietness: The impact of noise on Geneva rents. Urban Studies 42:633-646

Borrielo G, Chalmers M, LaMarca A, Nixon P (2005) Delivering real world ubiquitous location systems. Communications of the ACM 48:36-48

Brüel and Kjaer (2013) ANOMS - Airport Noise & Operations Monitoring System. http://www.bksv.com/products/environmentmanagementsolutions/airportenvironmentmanagement/anoms.aspx. Accessed 31 Dec 2013

Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D (2011) Gamification: using game-design elements in non-gaming contexts. In: CHI '11 Extended Abstracts on Human Factors in Computing Systems, pp 2425-2428, ACM, Vancouver

D'Hondt E, Stevens M, Jacobs A (2013) Participatory noise mapping works! Pervasive and Mobile Computing 9(5):681-694. Doi:10.1016/j.pmcj.2012.09.002

European Union (2002) Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise - Declaration by the Commission in the Conciliation Committee on the Directive relating to the assessment and management of environmental noise

IEC 61672-1:2002 (2002) Electroacoustics - Sound level meters - Part 1: Specifications

Kanjo E (2010) NoiseSPY: A Real-Time Mobile Phone Platform for Urban Noise Monitoring and Mapping. Mobile Networks and Applications 15(4):562-574. doi:10.1007/s11036-009-0217-y

Manvell D, Ballarin Marcos L, Stapelfeldt H, Sanz R (2004) SADMAM - Combining measurements and calculations to map noise in Madrid. Paper presented at the 33rd International Congress and Exposition on Noise Control Engineering (Inter-Noise 2004), Prague, 22-25 August 2004

Martí I, Rodríguez L, Benedito M (2012) Mobile application for noise pollution monitoring through gamification techniques. In: Proceedings of 11th International Conference on Entertainment Computing (ICEC 2012), Bremen, 26-29 September 2012. Doi: 10.1007/978-3-642-33542-6_74

McDonald P, Geraghty D, Humphreys I, Farrel S (2008) Assessing environmental impact of transport noise with wireless sensor networks. In: National Research Council (ed) Environment and Energy 2008. Transportation Research Record, is 2058. Transportation Research Board, Washington, pp 133-139

Modsching M, Kramer R, Hagen K (2006) Field trial on GPS accuracy in a medium size city: The influence of build-up. Paper presented at the 3rd Workshop on Positioning, Navigation and Communication, Hannover, 16 March 2006

Nüst D (2013) OpenNoiseMap. https://wiki.52north.org/bin/view/SensorWeb/OpenNoiseMap Accessed 31 Dec 2013

Ogneva-Himmelberger Y, Cooperman B (2010) Spatio-temporal Analysis of Noise Pollution near Boston Logan Airport. Urban Studies 47:169-182

Rekimoto J, Miyaki T, Ishizawa T (2006) LifeTag: WiFI-based continuous logging for life pattern analysis. In: Hutchison D et al. [ed.] Location and Context Awareness. Springer, Berlin

Stevens M (2012) Community memories for sustainable societies - The case of environmental noise. Dissertation, Vrije Universiteit Brussel

Theakston F (2011) Burden of disease from environmental noise: Quantification of healthy life years lost in Europe. World Health Organisation, Copenhagen

Weber D (2013) A Review of 30 Sound (Noise) Measurement Apps http://www.safetyawakenings.com/safety-app-of-the-week-42/ Accessed 30 Aug 2013

World Health Organisation (2007) Night Noise Guidelines (NNGL) for Europe. http://ec.europa.eu/health/ph_projects/2003/action3/docs/2003_08_frep_en.pdf Accessed 5 Aug 2013

World Health Organisation (2013) Noise: Facts and Figures. http://www.euro.who.int/en/what-we-do/health-topics/environment-and-health/noise/facts-and-figures. Accessed 5 Aug 2013

Part V
Cartography in Education

Maps of Mars Compiled by Students at Lomonosov Moscow State University

Zhanna Rodionova, Aleksandr Berlyant, Juliya Brekhovskikh, Evgeniy Lazarev, Maria Lazareva and Olga Nosova[1]

Abstract The first map of the far side of the moon was made with the participation of Lomonosov Moscow University (Sternberg Astronomical Institute MSU) in 1961. These mapping technologies were then used in preparing the "Complete Map of the Moon" in 1967, and other maps. There is a special course in "The Mapping of Extraterrestrial Objects" for students of the Geographical Department of MSU. Some students have made maps of Mars as graduation theses and are still working on this theme at the Sternberg State Astronomical Institute. For example J.A. Ilukhina (Brekhovskikh) compiled a hypsometric map of Mars at a scale 1:26 000 000. This map was published in 2004 with an edition of 5 000 copies. The maps of the northern and southern hemispheres of Mars have also been compiled for the hypsometric globes of Mars. M.S. Shibanova (Lazareva) made hypsometric maps of Phobos and Deimos at a scale of 1:60 000, published in 2011 with an edition of 200 copies. Later she prepared a more detailed version of the Hypsometric Map of Mars with a new hypsometric scale. This chapter describes the compilation of these maps.

Z. Rodionova, E. Lazarev
Sternberg State Astronomical Institute, Lomonosov Moscow University,
Universitetsky pr. 13, Moscow, Russia, 119992
e-mail: marss8@mail.ru, zhecka@inbox.ru

A. Berlyant, E. Lazarev, M. Lazareva, O. Nosova
Lomonosov Moscow State University, Geographical Department,
Vorobjevu Goru 1, Moscow, Russia, 119992
e-mail: aberlyant@hotmail.com, zhecka@inbox.ru, lekanosova@gmail.com

J. Brekhovskikh
Space Research Institute,
Profsouznaya 88, Moscow, Russia
e-mail: juliailuhina@yandex.ru

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_23,

1 Introduction

The most valuable information about the bodies of the solar system, surpassing the volume, quality, resolution, and other parameters of data from telescopic observations, were received over more than fifty years of space flight. The moon, the closest celestial body to the Earth, was the first object of extraterrestrial mapping. Experience gained in the process of creating lunar maps served as a basis for planetary cartography. A special course, "The Mapping of Extraterrestrial Objects" for students of the geography department has been run at Lomonosov Moscow State University since 1989. The main topics discussed are: the surfaces of the moon, Mercury, Venus, Mars and its satellites, the satellites of the giant planets, the history of mapping methods and showing land form features in overview maps, the horizontal and vertical basis, projections, thematic mapping and the use of cartographic research to identify global patterns in the surface structures of these celestial bodies. Many students show a preference for the mapping of Mars. J.A. Iluhina, for example, wrote her graduation thesis "Automatic Compilation of A Hypsometric Map of Mars" in 2001. She made a map and created a 3D model of Valles Marineris. Later she prepared a new version of the Hypsometric Map of Mars.

2 Hypsometric Map of Mars

A hypsometric map gives a geometrically accurate representation of a planet's relief with contour lines and colour-coded altitude levels. The Hypsometric Map of Mars (Figure 23.1), at scale1:26000 000 (Ilukhina and Rodionova 2004), was compiled at the Sternberg State Astronomical Institute (SAI) in cooperation with the Department of Cartography and Geoinformatics Faculty of Geography Moscow State University. The Mars Orbiter Laser Altimeter (MOLA) data was used for mapping at a resolution of 1° (64 800 points). A digital model of the relief was constructed using ArcGIS software, using spline interpolation for construction of the digital model. The map shows only one element – the relief – with the names of significant features labelled: terras, plateaus, mountains, lowland plains and some large craters.

Fig. 23.1 Hypsometric Map of Mars, scale 1:26 000 000. Compiled by Iluhina J.A., Editor Rodionova Zh.F.

Names on the map are given in both Latin and Russian versions. We used data published by the International Astronomical Union [http://planetarynames.wr.usgs.gov] and the catalogue by Burba (1981). The MOLA instrument on board the US Mars Global Surveyor spacecraft is a high-accuracy laser altimeter, which provided absolutely new information about Mars's surface from over 670 million individual measurements. The global topography of Mars is now known to a greater accuracy than Earth's continents in a root mean square sense. The laser spot provides elevations relating to an area of 160 x 160 meters. The aggregated data have a 1-2 km spatial resolution and a 1m height-precision (Smith et al. 2000). This data has provided more information about the planet's shape, its geodetic parameters, and they have been used for surface modelling. Elevation is reckoned from an equipotential surface known as the areoid, measured relative to a triaxial ellipsoid.

The height scale on the Hypsometric Map of Mars contains 21 heightsteps. Up to a height of 8 km the contour interval is1km; from 8 to12 km the interval is 2 km; and above 12 km the interval is 10 km. Most of the northern hemisphere of Mars is occupied by relatively smooth plains. Vastitas Borealis has a depth of - 4 to - 5 km as do Utopia Planitia and Acidalia Planitia. Arcadia, Chryse, and Amazonis Planitia are higher by 1 km. In the southern hemisphere the plains are relatively few. There are Hellas Planitia, 2300 km in diameter, with a depth of - 8 km and 800km in diameter; and Argyre Planitia with a depth of about 3 km. The average height of the highlands of Mars are 3-4 km. Syria Planum is located at the plateau heights of 5-6 km, Sinai Planum at 3 to 5 km, and Solis Planum is at 3 to 4 km, Hesperia Planum and Syrtis Major Planum at 1 to 2 km. The largest mountain is at the equator – Tharsis Montes, with a diameter of about 6000km and a height of 9 km. Above it tower three extinct volcanoes: Ascraeus, Pavonis and Arsia Mons, located on the same line. The volcanoes have height of 14-18 km. The highest volcano on Mars is Olympus Mons. Its height is 21 km. In the equatorial zone of Mars is a giant system of faults with steep slopes - Valles Marineris. It has a maximum depth of 6 km and at the widest part is about 700 km across.

3 Hypsometric Globes of Mars

The history of the creation of the Martian globes covers more than 150 years (Berlyant 2007, Blunck 1993, Hargitai and Gede 2009, Mokre 2005, Rodionova, 1991). Maps of the northern and southern hemispheres were created for the Hypsometric globes of Mars at Sternberg Astronomical Institute and the Space Research Institute by J.A. Brekhovskikh (J.A. Iluhina), edited by Zh.F Rodionova.

Fig. 23.2 Hypsometric globes of Mars in diameters 15 and 21 centimetres

The Hypsometric globes of Mars (Figure 23.2) are based on our hypsometric map of Mars, scale1:26000 000 (Ilukhina 2004) in diameters 15 and 21cm. The technology used to create the globe using thermoplastic materials allows for forming the hemispheres of the sheets with the printed maps gluing the hemispheres at the equator. The original hemisphere prepared with the image is placed in the plastic moulding device, and the hemispheres are pressed at high temperature using a metal template. Western and eastern hemispheres were recut as the northern and southern hemispheres in the azimuthal projection, taking into account the law of deformation occurring during forming of the plane in the hemisphere. Transforming the original cartographic image on the plane into an undistorted image on the field in this way can only be done if the distortion of the lengths of the meridians in the original projections will be constant and equal tom = 2/π, and the distortion of the lengths of the parallels of latitude is a function of the form n = (π-2φ):(π cosφ) (Boginsky et al. 1990). This requirement corresponds to an azimuthal equidistant projection to the plane of section passing through the centre of Mars. The projection is an orthogonal grid. The parallels are represented by equally spaced concentric circles, and meridians (the straight lines emanating from the centre of the circles). When placing the original map, we took into account of the fact that when forming the hemispheres the original image is stretched by more than half. As the tension is not even, the names of relief forms were arranged parallel to the equator. To construct the cartographic imageArcGis10 software was used. The final design was made in the Coral Draw graphics editor. Based on project data previously made in ArcGis, hypsometric maps were made for the southern and northern hemispheres and redesigned in the azimuthal projection, prepared by the grid. The map images in Eps format were then transferred to CoralDraw, where the labels were made for landforms, map grids and the layout for printing was prepared in accordance with the requirements of the publisher. The landing sites of spacecraft on the surface of Mars are also shown on the globes.

4 Hypsometric Maps of Phobos and Deimos

The history of investigations of the Martian moons Phobos and Deimos by spacecraft is very well described in "The International Atlas of Mars Exploration. The First Five Decades: 1953-2003" (Stooke 2012). There are several digital terrain models of Martian satellites based on different shaped models of Phobos (Turner 1978, Thomas 1989, Duxbury 1991, Simonelli et al. 1993, Willner et al. 2010, Zubarev et al. 2012) and Deimos (Duxbury & Callahan 1989, Thomas 1989, Thomas 1993). Note that only the most up-to-date models of Phobos were used to generate DTMs. Older Phobos, and all Deimos models, are shaped ones and contain merely a few hundred control points. For our maps and 3D-models we used digital elevation models (DEM) with 0.5 degree resolution (~100 m/pixel) for Phobos (240 747 object points) (Willner et al. 2010) and 5 degrees (~400 m/pixel) for Deimos (2700 object points) (Thomas 1993). Use of these DTMs is based on a combination of several factors: regular coverage of the satellites' entire surface, sufficient accuracy and data accessibility.

For mapping we used biaxial ellipsoids although triaxial versions with radii of 13.0×11.4×9.1km for Phobos and 7.8×6.0×5.1km for Deimos are recommended by the IAU (Archinal et al. 2011). The heights on Phobos and Deimos are referenced to spheres of mean radii 11.1 and 5.77 km respectively (Willner et al. 2010). Using different vertical (sphere) and horizontal (2-axial map) references for cartographical representation was based on references of source DEM (sphere) and not the spherical form of Martian satellite hemispheres referenced to biaxial ellipsoids. In mapping the surfaces of Martian satellites Phobos and Deimos, we can distinguish three particular cartographical factors: 1) their small size and irregular shape which influences the coordinate system and projection, 2) the typical relief features, which should be distinguished, 3) the difficulty of obtaining and processing remote sensing data (accuracy depends on it). Considering these issues we created hypsometric maps of the Martian satellites.

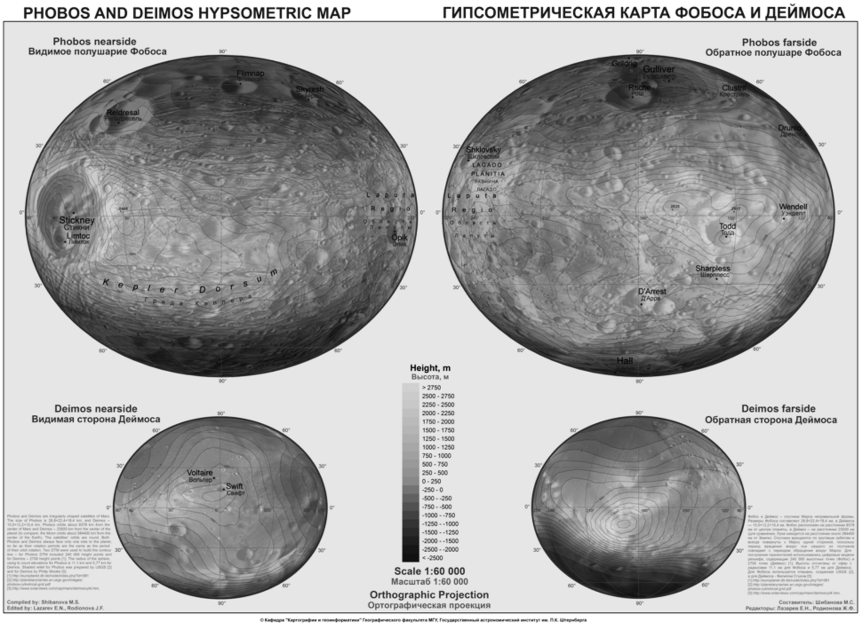


Fig. 23.3 Phobos and Deimos Hypsometric Map, scale 1: 60 000

The Phobos and Deimos relief map (Figure 23.3) was compiled in orthographic projection (Shibanova, 2011) and published in 2012 (Lazareva, 2012). Orthographic projection is a perspective projection with the centre in infinity, and therefore it is suitable for mapping celestial bodies (Bugaevsky and Tsvetkov 2000). It is important that in ArcGIS the orthographic projection works only for a sphere, so we used a local cartesian projection, similar to the orthographic projection, but also based on spheroids. Like the moon, both Phobos and Deimos always face the same side to the planet since their rotation periods are locked with the period of their orbital rotation. The central longitudes of the hemispheres are 0 and 180 degrees, corresponding to the satellites' nearside and farside respectively. The map grids were thus derived from the best-fit biaxial ellipsoids 11.2 × 9.1 km for Phobos and 6.0 × 5.1 km for Deimos. The special colour height scale with a fixed interval of 500m was designed with respect to the real colours of the satellites' surfaces. A scale of 1:60, 000 was selected for the map, which results in an acceptable resolution of 11.8 pixel/mm (300 dpi) for A1 format hard copy prints. All IAU approved Phobos and Deimos surface feature names (17 craters, Lagado Planitia, Laputa Regio and Kepler Dorsum on Phobos and 2 craters on Deimos) (Gazetteer 2013) are set both in Latin and in Russian variations. The map was compiled using the ESRI ArcGIS Desktop software.

Conclusion

A great deal of new information obtained by spacecraft about the surfaces of celestial bodies still requires processing and analysis. Students are interested in the surfaces of the moon, Mars, Venus, Mercury and other bodies. The special course "The Mapping of Extraterrestrial Objects" for students of the Geographical Department at Lomonosov Moscow State University prepares specialists in planetary cartography, and they have compiled Martian maps and globes for a wider public audience.

References:

Archinal B, A'Hearn M, Bowell E, Conrad A, Consolmagno G, Courtin R, Fukushima T, Hestroffer D, Hilton J, Krasinsky G, Neumann G, Oberst J, Seidelmann P, Stooke P, Tholen D, Thomas P, Williams I et al. (2011) Report of the IAU Working Group on Cartographic Coordinates and Rotational Elements: 2009. Celest. Mech. Dyn. Astron. 109:101-135.

Berlyant AM (2007) Globusy. GEOS. Moscow.

Blunck J, Jons HP (1993) Der Rote Planet Im Karten Bild. Staatsbibliothekzu. Berlin.

Boginsky VM., Rodionova ZhF, Myshetskaya EN (1990) Opyt sozdaniya globusa Luny na

termoplastichnom materiale. Geodeziya i Cartographiya., Moscow, № 6, p. 28-34.

Bugaevsky L, Tsvetkov V (2000) Geoinformatsionnie sistemy. Zlatoust, Moscow.

Burba GA (1981) Nomenklatura detaley reliefa Marsa. Nauka. Moscow.

Duxbury T (1991) Analytic model for the Phobos surface. Planet. Space. Sci. 39:355-376.

Duxbury T, Callahan J (1989) Phobos and Deimos control networks. Icarus 77:275-286.

Gazetteer of Planetary Nomenclature. http://planetarynames.wr.usgs.gov/ (Accessed 15 April 2013)

Hargitai H, Gede M (2009) Three virtual globes of Mars: topographic, albedo and a historic globe. EPSC Abstracts, Vol. 4, EPSC2009-47.

Ilukhina JA, Rodionova ZhF (2004) Gipsometricheskaya karta Marsa. 1:26 000 000. Moscow, Nauka.

Lazareva M (2012) Kartografirovenie reliefa sputnikov Marsa. LAP Lambert Academic Publishing, Saarbrucken.

Mokre J (2005) The Globe Museum of the Austrian National Library. Bibliophile Edition. Vienna.

Rodionova ZhF (1991) Space achievements and maps and globes of planets. Geodesy and Cartography, Moscow, N 7, p. 15-22

Shibanova M, Lazarev E (2011) Mars moons relief mapping and modelling: problems and solutions. EGU General Assembly 2011 Geophysical Research Abstracts 13:6822.

Simonelli D, Thomas P, Carcich B, Veverka J (1993) The generation and use of numerical shape models for irregular solar system objects. Icarus 103:49-61.

Smith DE, Zuber MT, Frey HV et al (2000). Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars. Journal of Geophysical Research, August 16, 1-38

Stooke P (2012) The International Atlas of Mars Exploration. The First Five Decades: 1953-2003, Cambridge.

Thomas P (1989) The shapes of small satellites. Icarus 77:248-274.

Thomas P (1993) Gravity, tides, and topography on small satellites and asteroids: application to surface features of the Martian satellites. Icarus 105:326-344.

Turner R (1978) A model of Phobos. Icarus 33:116-140.

Willner K, Oberst J, Hussmann H, Giese B, Hoffmann H, Matz K-D, Roatsch T, Duxbury T (2010) Phobos control point network, rotation, and shape. Earth Planet Sc. Lett. 294:541-546.

Zubarev A, Nadezhdina I, Patraty V, Shishkina L, Zharov O, Zharov A, Oberst J (2012) Phobos Control Point Network and Librations. EPSC Abstracts 7: EPSC2012-238-2

The Use of Cartograms in School Cartography

José Jesús Reyes Nuñez[1]

Abstract Cartograms are one of the youngest methods of representation in thematic cartography. Their origins can be dated to the first years of the 20th century, but the first cartogram-like representations were made in the second half of the 19th century. A very interesting detail is that the first cartograms were created to be specifically published in newspapers or journals destined for the public in general, as well as for publications related to school cartography. During the second half of the 20th century the development of cartograms decreased and their use in publications related to education can be considered all but nil, for various reasons that are analysed in this study. In the last 25 years we have witnessed the rebirth of cartograms, in many cases using the latest web-based interactive technologies for their publication. This chapter also provides recommendations for the possible use of cartograms to illustrate and support teachers' explanations in classrooms.

1 Brief introduction to cartograms

Various definitions of cartograms were penned during the 20th century and all agree on a common characteristic: the intentional distortion of the areas represented on an initial (base) map as a tool to represent a theme. Erwin Raisz – the first internationally recognised cartographer to develop continuous research on this theme – defined them simply as "diagrammatic maps" (Raisz 1962) and he also introduced the concept of "value-area" or "value-by-area" cartograms (Raisz 1948, 1962) to identify rectangular statistical cartograms. Norman J. W. Thrower (2008) defined the cartogram as "an abstracted or simplified map for displaying quantitative data for which the base is normally not true to scale", while according Slocum et al. (2005) it is a representation where "spatial geometry is distorted to reflect a theme".

Cartograms became popular in France in the 19th century and in the United States beginning in the 20th century, being published in newspapers or journals to illustrate and make articles written for the public more understandable in general, as well as those in textbooks and atlases.

J.J. Reyes Nuñez
Department of Cartography and Geoinformatics, Eotvos Lorand University,
Pazmany Peter setany 1/A. Budapest 1117, Hungary
Tel.: 36 1 372 29 75 Fax: 36 1 372 29 51
e-mail: jesusreyes@caesar.elte.hu, jesus@map.elte.hu

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_24,

In Germany, Professor Hermann Haack and H. Wiechel "published a cartogram depicting election results from the German Reichstag in 1903" (Fabrikant 2003), using rectangles and re-sizing them based on the population. One of the first cartograms published in the United States was the "Apportionment Map of the United States", made by Professor William B. Bailey (Yale University) scaling the size of the states to their population in 1910 and published in "The Independent" newspaper on 6 April, 1911 (Krygier 2008). In 1931 an English electrical draughtsman, Harry Beck, created his London Underground Map, reproduced in 1933 (Gießmann 2013). This was the first tube map conceived as a line-based cartogram: the starting point of a very popular and widespread method of representation for urban transport. Erwin Raisz researched this topic from 1934, constructing his first "value-by-area" cartograms, a representation of the world based on the use of rectangles re-sized according to the data and placed in a relatively similar position as the implied territories on the original map (Raisz 1948). The use of cartograms in different types of publications (including those for the public in general, as newspapers or journals) became popular in the United States in the first half of the 20th century. A singular example is the "sketch map" entitled "The World in Terms of General Motors" published in Fortune magazine in 1938 (Figure 24.1), in which the author (Richard E. Harrison) distorted the original projection to emphasise those territories (mainly North America, and to a lesser extent Europe) where the General Motors factories were located.

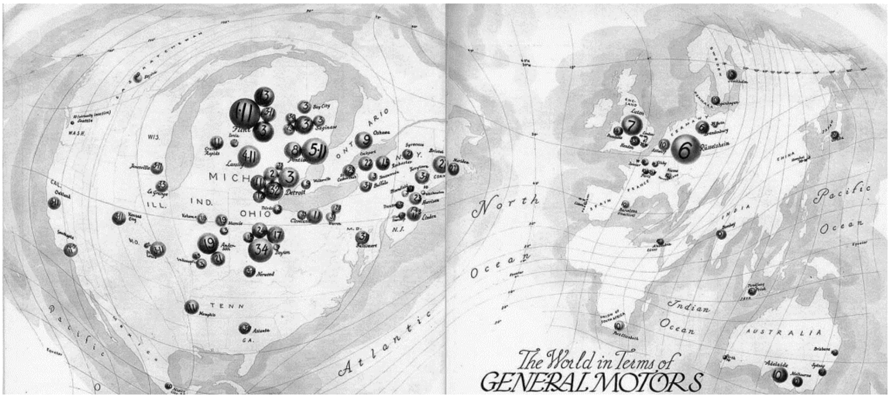


Fig. 24.1 Fragment of the map entitled "The World in Terms of General Motors"

More methods to make cartograms were developed in the 20th century, but I make mention of only one created in the last decade of the past century, which at present is internationally acknowledged: the Dorling or circular cartograms created by Daniel Dorling (1991). Beginning in the 21st century new solutions were developed for area cartograms: one is the method of building rectangular cartograms, as defined by Bettina Speckmann, whose principle was to resize and move each rectangle (symbolising, for example, a country) on a rectangular map layout according the values to be represented (Florisson et al. 2005). Michael Gastner and Mark

Newman created the diffusion-based method for density-equalising maps (Gastner & Newman 2004), which is also used by Dorling in his very popular "Worldmapper" project (Dorling et al. 2006).

In the present work I used the term "cartogram-like representations" (predecessors of the current cartograms), when the original geographic or political areas are substituted by geometrical shapes and the areas of these shapes are modified to represent the selected data. The placement of the shapes is arbitrary, very often without any relation or reference to their real geographic location.

Two denominations were also introduced to differentiate area cartograms for school cartography. Geographic area cartograms are those in which the original geographic unit or area is distorted according to the data, trying to keep their geographic location and adjacency with neighbours at least approximately. The same conditions should be fulfilled when creating a geometric area cartogram, but only after substituting the original geographic areas with a geometric shape (rectangle, circle, etc.).

2 Early cartograms related to school cartography

The concept of distortion in the cartograms made in the 19th century implied the substitution of the original geographic areas with geometric shapes (more often a circle or rectangle) rather than modifying the geographic area itself. The first cartogram-like representation using rectangles showed European countries according to their sizes in 1870, which was made by the French economist, geographer and educator Pierre Émile Levasseur (1875). He is undoubtedly the pioneer of the use of cartogram-like representations in school textbooks: other (more simple) examples can be found on page 778 of his geographic textbook, "La France, avec ses Colonies..." published in Paris in 1875 (Figure 24.2).

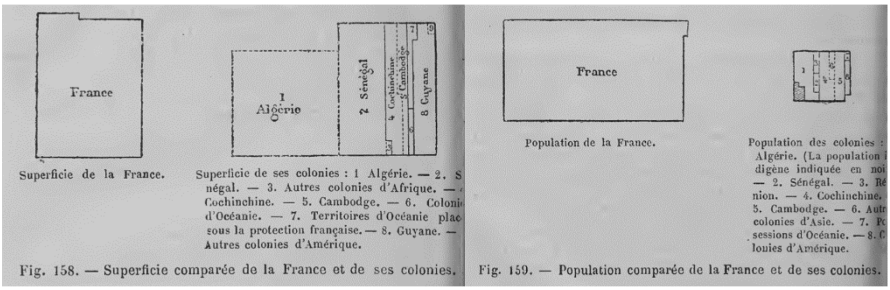


Fig. 24.2 Cartogram-like representations comparing the territorial extension and population of France and its colonies, published by Levasseur in 1875

Cartogram-like representations can be found in atlases for schools that were published in the United States of America during the 19th century. One of the earliest examples dates from 1837, when William C. Woodbridge published a "Chart of the comparative magnitudes of countries…" in his "Modern atlas, on a new plan, to accompany the system of Universal Geography", which was also completed with a list of questions to be put by the teacher to pupils in the interest of helping the analysis of data (Figure 24.3). This "comparative chart" was improved and reproduced in the Woodbridge's "Modern Atlas" of 1843 and in the "School Atlas, to accompany Modern School Geography" of 1845.

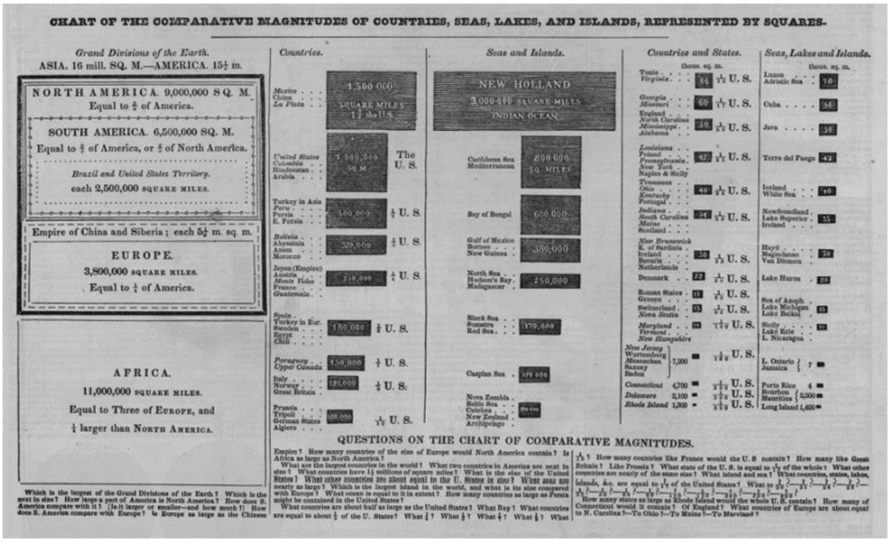


Fig. 24.3 Cartogram-like representation made by Woodbridge in 1837 (left side of the page)

In 1837 Jesse Olney also presented a coloured cartogram-like representation in his "New and Improved School Atlas", "exhibiting the comparative size, population, form of government and number of square miles in each of the principal empires, kingdoms, &c. of the globe" (Figure 24.4).

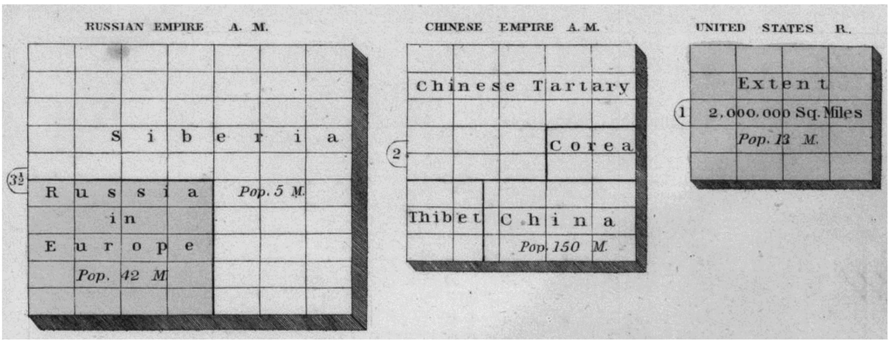


Fig. 24.4 Fragment of the cartogram-like representation made by Olney in 1837

The Rand McNally World Atlas of 1897 also included some cartogram-like representations named “statistical diagrams”, one of them representing countries with two circles, the diameter of which varied according to the area and the population in 1890 (Figure 24.5).

Fig. 24.5 Fragment of statistical diagrams in the Rand McNally World Atlas of 1897

However, Erwin Raisz can be considered the first cartographer to develop scientifically based research on cartograms, publishing his results for the first time in 1934. In his article entitled “The Rectangular Statistical Cartogram” he described the method used to create this map-derived representation, specifying that “*the statistical cartogram is not a map. Although it has roughly the proportions of the country and retains as far as possible the relative locations of the various regions, the cartogram is purely a geometrical design to visualize certain statistical facts and to work out certain problems of distribution*”. He also described the educational role of a cartogram: “*Its educational value is not limited to the schools: it may serve to set right common misconceptions held by even well informed people…*”

Raisz was also the first professor and researcher who consistently dedicated a chapter to cartograms in his textbooks on cartography, beginning from the first one published in 1938. One of the characteristics of these chapters is that Raisz not only included notions about the value-by-area cartograms, but also presented less used methods of representations based on maps, such as traffic-flow maps, centrograms, etc. In his textbooks Raisz always put emphasis on the role that cartograms can play in geographic thinking. In 1938 he wrote: “*Cartograms are useful tools of modern geography, and their possibilities are not yet fully explored…Value-area cartograms help a great deal in our geographic thinking… These cartograms also form a useful base for showing ratios.*” (Raisz 1938). In his “Principles of Cartography” (1962) we can also read: “*Value-by-area cartograms are important. Our socioeconomic overview of the world will be more realistic if* *we think of the relative importance of its parts in the proportion of a population cartogram rather than in the proportions of a map. The results are often quite surprising. Such cartograms can make certain problems startlingly clear…*". Furthermore, in the article of 1934 he also raised an interesting idea that was never applied in practice: "*If natural geographic regions could be used instead* [of political boundaries], *the cartograms would be still more instructive.*"

Another interesting detail is that in these chapters Raisz also included explanations related to the "diagrammatic maps in teaching" (Raisz 1938) or "blackboard maps" (Raisz 1962). Raisz understood the simplified outline maps that teachers often sketch on the blackboards in these terms, arguing that "*students understand them better than wall maps with all their complexities*".

I have developed personal research on the use of cartograms in American school atlases in the 20th century. First I examined original atlases in the Department of Cartography and Geoinformatics library in Budapest and completed the research with atlases that can be found on the internet. My research shows that cartograms were not used significantly in American school atlases during the 20th century. Area cartograms can be found in some general atlases that were also used in education, published in North America in this period: a good example is the Atlas of Global Geography edited by Raisz in 1944, in which value-area cartograms were used to present data related to population, average incomes, death rates, nutrition and illiteracy. He also used a kind of cartogram-like representation – which can be better considered "diagrammatic maps" – to show the possible distribution and resizing of lands based on then theories of world geopolitics (Figure 24.6).

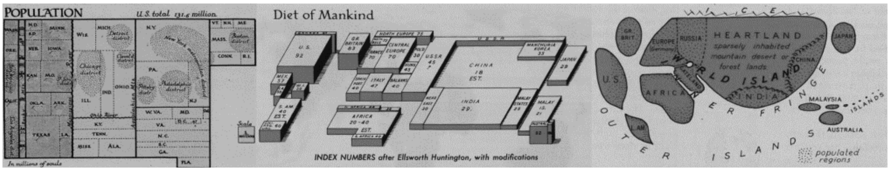


Fig. 24.6 Examples of value-area cartograms and diagrammatic maps in the Atlas of Global Geography. From left to right: population by state in USA (p. 44), nutrition standards by countries (p. 55) and the world according to the theory formulated by MacKinder in 1904 (p. 57)

Thirty five years later the Atlas of Canada and the World was published (Philip 1979), using rectangular area cartograms to represent topics such as population, import and export, and other area cartograms to show the world trends in oil production and consumption (Figure 24.7).

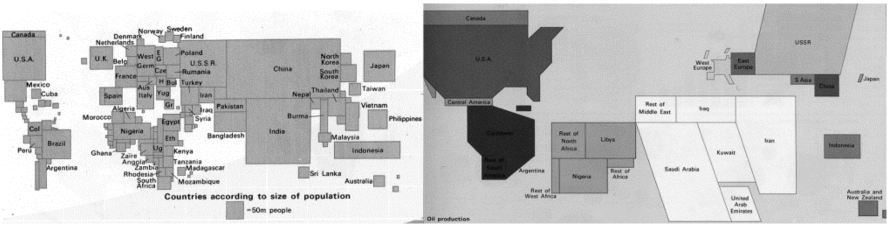


Fig. 24.7 Examples of cartograms in the Atlas of Canada and the World. Left: the world according to the population by countries. Right: oil production in the world

Some of the most important and internationally recognised European cartographic firms also included cartograms or cartogram-like representations in their school and general atlases. Among them should be mentioned the series of school atlases produced by the German Westermann firm (Westermann 1996) including area cartograms on the inside back cover (Figure 24.8), or by the Austrian Freytag-Berndt firm, using linear cartograms to present public transport in Vienna (Freytag-Berndt 1998). More recently, pupils can find cartograms in the Oxford International Student's Atlas (Wiegand 2006b), which was published by Oxford University Press in 2006 (Figure 24.9).

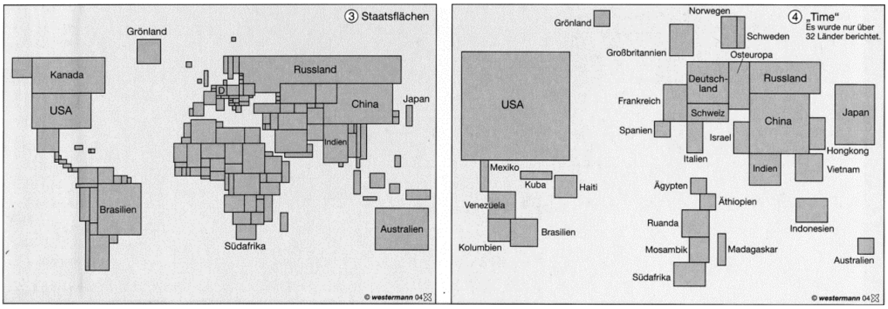


Fig. 24.8 Rectangular cartograms in the Westermann School Atlas (1996)

The use of cartograms in general decreased during the second half of the 20th century until the decade of the 1990s. From this decade to the present we have witnessed a revival of the use of cartograms as a method of representation that has been adapted to the new graphic solutions for the geovisualisation of data on a powerful new media: the World Wide Web.

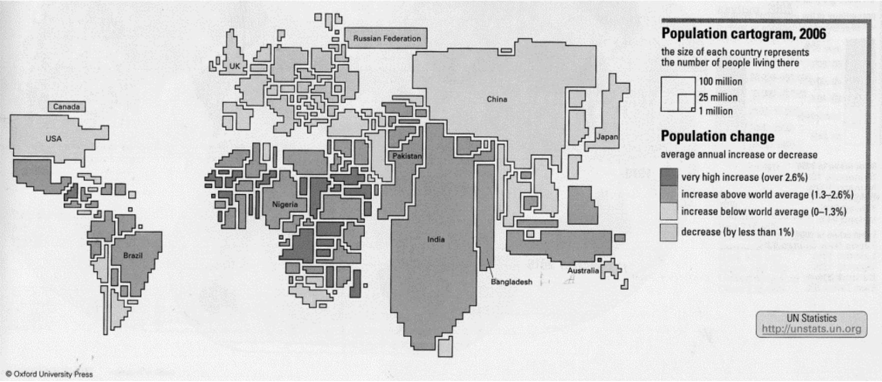


Fig. 24.9 Population-based cartogram combined with colour fill to represent two related topics: the size of each country represents the number of people living there, while the fill represents the annual increase or decrease of population (Wiegand 2006b)

3 How could cartograms be used in current school cartography?

Patrick Wiegand confirmed in Chapter 12, "Practical map activities: age 11 to 14" of his "Learning and Teaching with Maps" book (2006a) that when developing the map and atlas skills of pupils "*they should be able to draw conclusions from cartograms (for example, on a world population cartogram, where the size of the countries is shown proportional not to area but their population, identify countries that have a small area and large population)*". Earlier in the same chapter he also stated that "*they should begin to be competent in the use of rail and bus maps (including transformed maps)*", alluding to the line cartograms widely used in public transport.

Internationally recognised institutions like the Royal Geographic Society, together with the Institute of British Geographers, also recommend the use of cartograms in their tutorials on different themes, such as Population and Migration at KS3 level (Key Stage 3, grades 7 to 9, age 11 to 14): "*As well as the map of world population densities, students can look at proportional cartograms such as are available on the interactive Worldmapper website.*" (RGS-IBG 2013).

Compared with other fields of cartography, there is an as yet subdued, modest interest in using cartograms in primary and secondary (K12) education, mainly at the secondary level. Nevertheless, different materials dedicated to this theme can be found on the internet, such as a lesson planned within the curricular material entitled "Resources for Teaching About the Americas (RETANet)", which was designed by the Latin American and Iberian Institute at the University of New Mexico (Pelkofer 1996). Some authors in other countries also recommend their use in teaching activities: Y. K. Singh (India) specified that "*converting recorded* *data into graphical tools for better and clearer understanding… interpreting graphs, charts and cartograms*" should be one of the assignments on making and handling tools to be given by the teachers for the pupils (2007).

Despite the examples mentioned above, cartograms did not become a commonplace tool when teaching geography in schools. In the interest of finding the cause of this situation we should examine those geographic topics where cartograms can help teaching activities, as well as the negative and positive aspects of using a cartogram.

3.1 Themes that can be represented by cartograms

When determining which topics can be represented with an area cartogram, I adhere to the opinion expressed by Daniel Dorling (1996), that area cartograms can be more effectively used to emphasise human geography than topographical representation. This means that cartograms are a powerful tool to make more visible those geographically smaller units or areas that are significant within a selected theme. The most traditional examples are the cities or metropolises, which on a small-scale map occupy a very small area, when their population can be the largest and densest in a given country. The small size of the original area on the map can mean that an inexperienced map user (in our case a pupil) gives less importance to a city than it really has. There are a considerable number of themes to be represented, because human geography studies a wide spectrum of activities related to society, such as population, economy, development, health, culture, etc.

3.2 Difficulties and solutions using cartograms in school

Obviously, distortions make the trends between data more easily identifiable visually, but also can make the represented geographic units or territories unrecognisable (or much more difficult to recognise). This fact can notably limit the use of cartograms not only in school cartography, but in general. The distortion of territories (including their substitution and resizing with geometric shapes) can also provoke a different placement of the territorial units in relation to their original geographic position. In geometric area cartograms the displacement of original units is an unavoidable characteristic, and it becomes more significant when using the Dorling cartogram, when geographic or political units are substituted with tangent circles. The best applicable solution to this situation is using cartograms only to represent territories well-known by pupils (the traditional position of continents on a world map, their own country or own continent). This should be considered the most important premise, a necessary initial condition for their use in school cartography. Cartographers can also strive to determine the distortion more carefully, making the differences between units visible and trying to keep their real location as accurate as possible, but it is very difficult to accomplish this condition in practice.

A negative characteristic of cartograms is that reading exact information about the represented data is more difficult than on thematic maps made with other methods of representation. The main task of cartograms is not to give exact information about the data, because area cartograms are created for faster visual identification of the trends between units of a determined topic. The legend made for some geographic area cartograms can be difficult for a pupil to understand, but a geometric area cartogram can be accompanied by a clear and simple legend that makes possible the approximated determination of the represented value and the comparison between the different sized geometric shapes. The most modern technologies also offer us interactive tools that can substitute for the traditional legend of a map or can complete the data represented on the map with new information.One of them is the use of "call-outs" on interactive maps presented on the internet has gradually become a tradition. A call-out is a label framed within a balloon, which appears after clicking on an area of the map. In this way we can give the exact amount of data represented by the cartogram and if needed, other information related to the current theme (Figure 24.10).

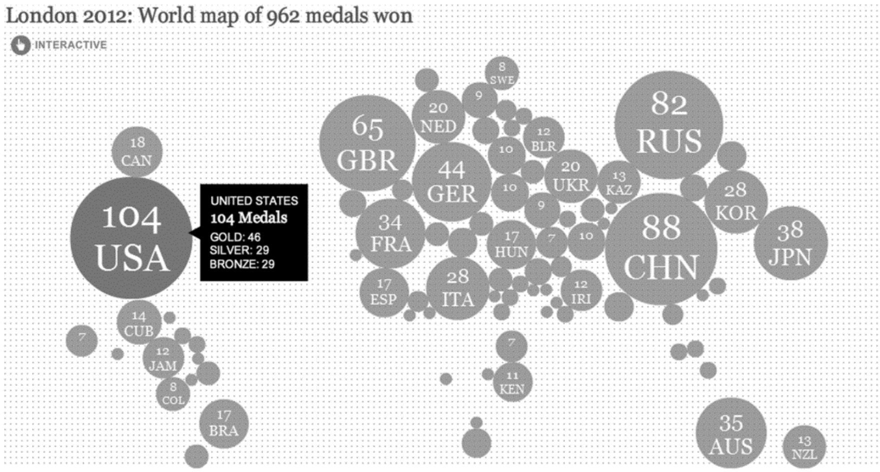


Fig. 24.10 Internet-based interactive Dorling cartogram presenting results of the Olympic Games in London (2012) <http://www.telegraph.co.uk/sport/olympics/9436640/London-2012-Olympics-dynamic-world-medal-map.html>

3.3 Preliminary suggestions

The use of cartograms in teaching activities should never be intended to replace traditional thematic maps, but only to complete the information given by these maps. Cartograms have an important feature: they are able to break (in a figurative sense) the geographical and political borders, increasing or decreasing the size of the units to emphasise or obviate their importance within a given theme, helping in this way to make the role played by the unit in the selected theme more understandable.

Following on from the previous explanations, revitalising the use of area cartograms in school cartography can mean:

- Returning to the general principle applied during the making of the first cartograms and cartogram-like representations in the 19th century, further developed by Raisz in the 20th century. Some can be also characterised as "map-derived diagrams" (see Figure 24.4, 24.5 and 24.6), which were created to facilitate the visual comparison of units by converting them to resized geometric shapes. This is also a return to the re-valuation of the diagrammatic aspect of cartograms in a way similar to that followed by Raisz in his textbooks and articles.
- Introducing the use of the newest methods to create area cartograms (e.g. Gastner-Newmann, Dorling, Speckmann) supported by computer- and internet-based technologies.

4 Future research

The possible use of cartograms in school cartography, in the later grades of elementary schools and in secondary schools, is a theme that has not yet been researched. In recent years, only one study can be listed, related to the topic of using contiguous geographic area cartograms in education: it was developed by Kaspar et al. (2011) on the use of cartograms by 50 students from a cartography class at the University of Applied Sciences in Karlsruhe, who were 24 years old on average. They concluded that in general "*the more commonly used choropleth/graduated circle map combination yields more effective (accurate) and more efficient (faster) responses than the cartogram maps*". But at same time the authors stated that "*for simple tasks, cartograms seem as effective and efficient as the more traditional mapping method*", while "*for complex inference questions, inference performance with cartograms is significantly dependent on whether regular or irregular zones are distorted*".

Research should be developed on this theme. There is a "pre-disposed" negative attitude or generalised assumption that cartograms are very difficult to be read and understand for younger pupils, but it has never been empirically measured and proven. Four main questions should be tested in future studies:

- Which themes are recommended for representation on cartograms made for pupils?
- How can pupils identify the known geographic units after their distortion on cartograms?

- Are there any differences between the readability and understandability of geographic and geometric area cartograms by the pupils?
- How can the more and less traditional area cartograms be upgraded or completed using other methods of representation, as well as supported with the graphic solutions offered by the newest technologies? Can pupils understand these more complex representations?

Finding answers to these open questions should be a primary task for the specialists working on this theme not only to make the cartographic education of the new generations more multifaceted, but also to provide a scientifically solid base for its future development.

References

Dorling D (1991) The Visualization of Spatial Structure. Dissertation, University of Newcastle upon Tyne.

Dorling D (1996) Area cartograms: their use and creation (CATMOG 59).http://qmrg.org.uk/files/2008/11/59-area-cartograms.pdf. Accessed 12 December 2013

Dorling D, Barford A and Newman M (2006) Worldmapper: the world as you've never seen it before. http://www.dannydorling.org/wp-content/files/dannydorling_publication_id0892.pdf. Accessed 18 February 2014

Fabrikant S-I (2003). Commentary on "A History of Twentieth-Century American Academic Cartography" by Robert Mc Master and Susanna McMaster. *Cartography and Geographic Information Science (CAGIS)*, 30(1), pp. 81-84.

Freytag-Berndt (1998) Unterstufen-Schulatlas. Freytag-Berndt, Wien

Florisson S, van Kreveld M. and Speckmann B (2005) Rectangular Cartograms: Construction & Animation.http://www.win.tue.nl/~speckman/papers/CartoAnim.pdf.Accessed 20 December 2013

Gastner M-T and Newman M-E-J (2004) Diffusion-based method for producing density-equalizing maps. In: Proceedings of the National Academy of Sciences of the United States of America (PNAS), 101(20). Available via PNAS website. http://www.pnas.org/content/101/20/7499.full.pdf. Accessed 12 December 2013

Gießmann S (2013) Henry Charles Beck, material culture and the London Tube Map of 1933. http://amodern.net/article/henry-c-beck-material-culture-and-the-london-tube-map-of-1933. Accessed 18 February 2014

Harrison R-E (1938) The world in terms of General Motors. In: Fortune magazine [online]. http://www.fulltable.com/vts/f/fortune/aac/0d5.jpg. Accessed 20 December 2013

Kaspar S, Fabrikant S-I and Freckmann P (2011) Empirical study of cartograms. http://www.geo.uzh.ch/~sara/pubs/kaspar_fabrikant_freckmann_ica11.pdf. Accessed 20 December 2013

Krygier J (2008)1911 Cartogram: "Apportionment Map". Available via Making Maps: DIY Cartography website. http://makingmaps.net/2008/02/19/1911-cartogram-apportionment-map/. Accessed 11 December 2013

Levasseur P-É (1875) La France, avec ses Colonies. Contenant 174 cartes, coupes et figures explicativesinséréesdans le texte. https://archive.org/details/lafranceavecsesc00leva. Accessed 12 December 2013

McNally R (ed) (1897) Indexed Atlas of the World. Rand McNally& Co, Chicago

Olney J (1837) A new and improved school atlas, to accompany the practical system of modern geography. Robinson, Pratt & Co, New York

Pelkofer T (1996) Using cartograms to learn about Latin American demographics. http://laii.unm.edu/outreach/common/retanet/science_math/RETANET_Lesson-Plan_Using-Cartograms.pdf. Accessed 15 December 2013

Philip G. (ed) (1979) Atlas of Canada and the World. George Philip and Son, Milwaukee

Raisz E (1934) The rectangular statistical cartogram. Geographical Review 24(2):292-296

Raisz E (1944) Atlas of Global Geography. Global Press Corp, New York

Raisz E (1948) General Cartography. 2nd ed. McGraw-Hill, New York

Raisz E (1962) Principles of Cartography. McGraw-Hill, New York

RGS-IBG (2013) KS3 Tutorial: Population and Migration. http://www.rgs.org/NR/rdonlyres/F7B60E25-E130-4E7E-B21B-6A32B25740C5/0/KS3_CPD_PopulationandMigration_script.pdf. Accessed 20 December 2013

Singh Y-K (2007) Instructional technology in education. APH Publishing Corporation, New Delhi

Slocum T-A, McMaster R-B, Kessler F-C and Howard H-H (2005) Thematic Cartography and Geographic Visualization. 2nd ed. Pearson Prentice Hall, Upper Saddle River

Thrower N-J-W (2008) Maps and Civilization: Cartography in Culture and Society. 3rd ed. The University of Chicago Press, Chicago

Westermann Kartographie (1996) Westermann Diercke Weltatlas. 4th ed. Westermann, Braunschweig

Wiegand P (2006a) Learning and teaching with maps. 1st ed. Routledge, New York

Wiegand P (ed) (2006b) Oxford International Student's Atlas. Oxford University Press, Oxford

Woodbridge W-C (1837) Modern atlas, on a new plan, to accompany the system of Universal Geography. Belknap & Hamersley, Hartford

Cartography as Intentional Distortion

Ferjan Ormeling[1]

Abstract Cartographers distort reality to allow map users to perceive it correctly. Hill-shading, classification, generalisation etc. are all methods that aim at simplifying reality and making it easier to comprehend. Similarly, by opting for specific mapping methods, reality is distorted, and specific data categories are highlighted. Map projections are another type of distortion, that lead map users on to perceive wrong relationships. All these types of distortion – aimed at a better understanding of reality - have in common that cartographers know about their side effects, and can thus offset them. As cartography has been democratized, everyone now is able to visualise spatial data, but is not necessarily aware of this inherent distortion. It should be the cartographers' task to make this knowledge available, and to indicate, for instance, wherever this might apply, for which type of map use and analysis the resulting maps would be (un)suitable.

1 Introduction

No map is the only true map. Maps are social constructs and depend on the views of their makers, each with their own preferential classification systems or colour schemes. Map design depends on the map's purpose, the visualisation objectives and the target audience; the contents of maps also depend on what spatial data society has collected and made available for mapping. The information policy of the government in many countries is still the decisive factor that determines which items can be mapped and which cannot, and this might lead to a one-sided view of reality even if this governmental information monopoly is increasingly undermined by GoogleEarth, OpenStreetMap and marketing firms that collect socio-economic data for postal code areas, all with their own agendas.

Cartographers have as their objective, space-related information transfer, and in order to communicate the essential aspect of their topics, they deliberately distort reality: they follow the motto of Paul Klee: "Art does not reproduce the visible; rather, it makes visible." (Klee 1920). I claim this same role for cartography. In order to make reality visible, cartographers have to distort it. We distort intentionally in order to provide a proper, better view of reality. The Swiss cartographer

Dr F.J. Ormeling
Faculty of Geosciences,
Utrecht University,
2 Heidelberglaan, 3584TS Utrecht, Netherlands
e-mail: f.j.ormeling@uu.nl

T. Bandrova et al. (eds.), *Thematic Cartography for the Society*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-08180-9_25,

Rudolf Knöpfli said in 1992 that maps do not render landscapes in more or less simplified form, they make visible specific characteristics of that landscape (Knöpfli 1992).

Sometimes however, this distortion also affects the truth – there are plenty examples of deliberate distortions of the truth resulting in maps that lead their viewers to have distorted views of reality.

Although these maps may have been correctly produced in the sense that they are repeatable and that the correct statistics have been used and visualised proportionally, they still might distort reality, due to the selection of a too small representative value of the symbols, for instance. Or they may be dishonest because some relevant information might conveniently not have been displayed. In a worse case, one might want to mask the location of military installations – in an earlier time these would have been left blank on the map, but as the very blankness of areas might attract unwelcome visitors, military areas might be displayed on the map as landscape parks. Maps might be dishonest because, instead of displaying reality, the information has been adapted to the message one wants to convey, for instance that one railway line is shorter and thus more efficient than another, by reducing the number of bends in its track.

One might object that only a very thin line divides this mala fide practice from the supposedly beneficial one first used by Harry Beck in 1931 for his London Underground network map (Garland 1994) which has become a landmark or icon for beneficial distortion: by straightening the course of the lines, reducing the angles and blowing up the map centre so that all station names could be rendered horizontally, with the same type size. It is a brilliant design, as it allows for individual clarity instead of overall clarity. Individual clarity is needed here in order to find out where exactly one has to change trains in order to get to a specific destination (Figure 25.1).

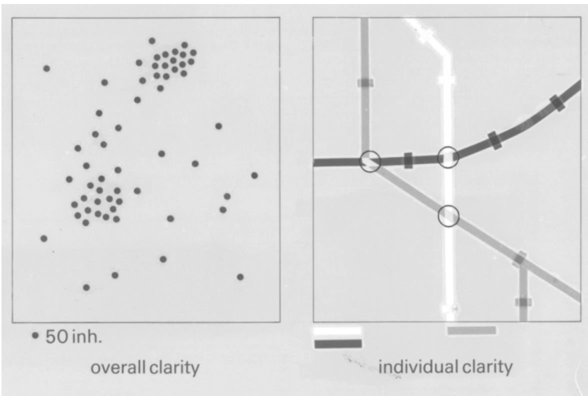


Fig. 25.1 Overall clarity vs. individual clarity

But cartographers should not close their eyes to the negative effects of this distortion. The same distance on the map between two consecutive tube stations may take one minute or 20 minutes, depending on whether one is within the city centre or in its outskirts. Stations that are located far apart on the map might in reality be situated on opposite sides of a street. Distortion was aimed here at maximum legibility and not at an optimal rendering of spatial relationships. So distortion comes at a cost, even if it is well-intended. It is cartographers who should be aware of these costs themselves and make their clients aware of them as well - this will be the central message of my presentation. These warnings to map users are a necessity, because maps often tend to be used for purposes other than they were intended for originally. The purpose of the intentional distortion of the Underground map was to effectuate individual clarity, so that map users could see how to get from one Underground station to another. Warnings will prevent them from using the map for other purposes, such as for travel time estimations, or making guesses about how far underground stations are apart in reality.

Sometimes we distort reality out of sheer ignorance, for instance when we don’t realise that by presenting spatial information in a specific sequence, a wrong impression might be created. In Figure 25.2, for instance, the school atlas spread suggests that Arabia is located east of India, a misconception that could have been easily prevented, by changing the sequence.

In this chapter I intend to look into different types of – beneficial – distortion and discuss their advantages and disadvantages, and will cover generalisation, relief representation, current thematic map types (like chorochromatic or choropleth maps), anamorphoses and map projections.

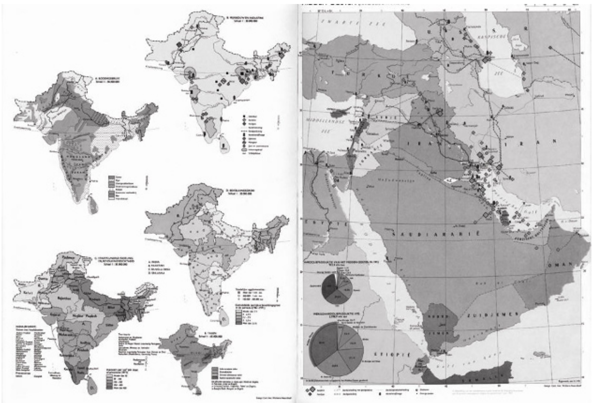


Fig. 25.2 Spread from a Dutch school atlas (Ormeling (ed.) Grote Bosatlas 43rd edition,1964)

2 Generalisation

As stated above, according to the Swiss cartographer Knöpfli, by generalisation we refer to the process of adaptation of a map’s contents so that it can show the meaning of the mapped area for a specific purpose (Knöpfli 1992). Knöpfli does not call this distortion, he regards it as an abstraction process, but the impact of this intervention can be quite drastic, leading to the relative neglect or even total obliteration of complete object categories from the data file.

Let us have a look at the Rhine valley, for example, in Figures 25.3 and 25.4. For a part of its course it runs through a plateau area (Rheinisches Schiefergebirge) and in order to prevent expensive construction works to bridge the side valleys, both roads and railroads are constructed within the Rhine valley itself, which makes this valley rather crowded. In order to be able to convey on smaller scale maps the notion that these roads and railways are running through the valley, the contour lines of the plateau slopes are displaced away from the river – a better example of distortion, when even hills are relocated, can hardly be found.

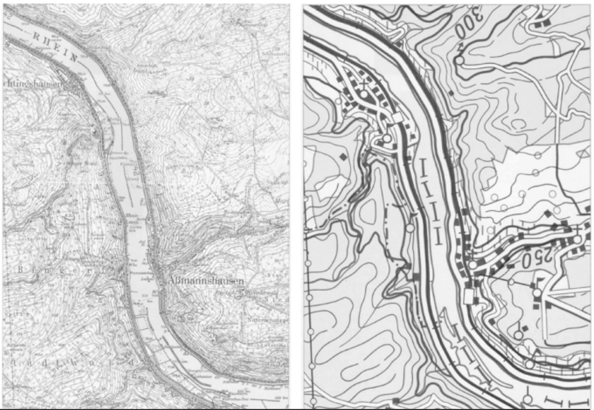


Fig. 25.3 Displacement of contour lines on a 1:200 000 map of the Rhine valley, generalised from a 1:25 000 map, visualised in maps (Kranz 1966).

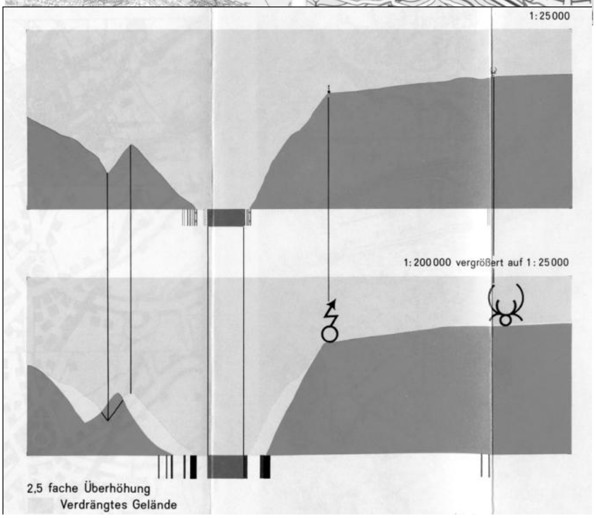


Fig. 25.4 Displacement of contour lines on a 1:200 000 map of the Rhine valley, generalised from a 1:25 000 map, visualised in cross-sections (Kranz 1966).

Again we should realise what has been sacrificed because of this distortion – the correct relative rendering of the road and rail network has been realised at the expense of a change in the absolute location of the plateau. The map has been made suitable for traffic analysts and road and rail users, but because of this distortion it is, for example, no longer of use to geomorphologists. Actually, users should be warned against the negative effects of this distortion for many map use tasks.

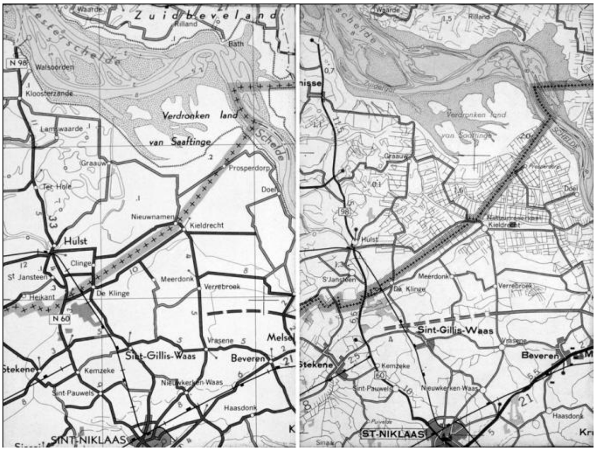


Fig. 25.5 and 25.6 Extracts from Dutch (left) and Belgian 1:250 000 topographic maps of the same border area.

Generalisation still remains subjective, even now, because different countries may decide to use different software programmes for generalisation. In the manual generalisation days, the same NATO generalisation rules applied to all maps in the 1:250 000 series produced in NATO countries, but Dutch and Belgian topographers still managed to come up with rather different images of the same area. As roadside canals were something special in Belgium, only occurring in coastal Flanders, these canals were retained on the Belgian 1:250 000 map. In the Netherlands, roadside canals and ditches are necessary almost everywhere, for drainage purposes, and that is why they are barely represented at smaller scales. Generalisation not only aims at simplifying the image for reduction to smaller scale, but also at characterising subregions, and that is why the Belgian map retained the many canals for coastal Flanders (see Figures 25.5 and 25.6).

Examples like these topographic maps, or of generalised town plans, show us that when we distort reality in order to serve one purpose, we lose out on other purposes. In most generalised town plans we can find our way easily: we immediately see the land use differences, but the maps became totally unsuitable for effectuating land use measurements because of the exaggeration inherent in generalisation. As can be seen when comparing it with the original, non-generalised map, the surface area of, for instance, the road network appears to have doubled or almost trebled in the generalised map.

3 Relief Representation

Panorama maps usually are good examples of relief distortion: in order to be able to perceive the relief characteristics, the vertical scale of these maps is exaggerated. In this way not only can the different relief characteristics be better differentiated, but the image has also been dramatised, and will thus be better remembered by the map viewers. Panorama map cartographers are also master distortionists, as they can bend the light – when looking in a distinct direction, they can line up mountains that actually extend themselves beyond that scope. Their practice is also to render, for example, the high Alpine areas relatively larger for their summer panoramas, and the lower mountain slopes, suitable for skiing, would be drawn relatively larger in their winter panoramas, thus adapting the landscape to the wishes of hotel owners and ski instructors. Another of their distortions or accomplishments, depending on one's views, was that they are able to render landscapes dreamlike, soft and grainy like the photographs of David Hamilton.

4 Thematic Map Types

4.1 Chorochromatic maps

This is one of the most used map types. In many cases, when the map subject is not earth-surface related, this map type can give a wrong impression, due to either the wrong colours or because of a false suggestion of quantitative differences, or both. Take world religion maps in conventional school atlases: when one wonders which group is larger, the green-coloured Muslims or the blue-coloured Hindus, the map would suggest Muslims, simply because the geographical area they inhabit is larger. Every reader knows that the green colour only stands for a local Muslim majority. Thus the map shows the location of areas where the majority of people are of the Muslim religion, and not how many people are involved. This would be the prevalent general reaction to the map, however – so it distorts! Even if it is correctly designed, a map that leads to the wrong reaction distorts reality. This has been recognised in some school atlases, such as the Diercke atlases (Westermann 2010), which therefore do correct this wrong reaction by adding the correct quantities with proportional symbols. One can then see that the number of Hindus is almost as large as that of Muslims; although the Muslim-dominated area is 10 times larger, it mainly consists of deserts. This distortion is caused by a false suggestion of homogeneity inherent in this map type. Something similar may happen in choropleth maps.

The use of chorochromatic maps for the display of nationalities or languages has been particularly harmful to Hungary in the past. Ethnographic maps produced in German-speaking countries favoured the use of prominent reddish colours for their own ethnic designation, and for Hungarians either a white or very light colour was used, which would earlier have been associated with the total absence of population rather than with a specific ethnic group. On the other hand, if this prominent red colour had been assigned to Hungarians, the image would be different (Krallert 1961).

4.2 Choropleth maps

In choropleth maps the observed data is usually categorised or classified and this procedure, aimed at simplifying the resulting image, of course also distorts: instead of individual observations, classes of observations are rendered by specific tints. That the resulting images are simpler and thus easier to read and to remember stands out when comparing them with class-less maps. In this process a loss of volume occurs, termed 'blanket of error' by the American cartographer George Jenks, which can be smaller or larger, depending on the classification methods and the number of classes that cartographers opt for (Jenks 1967). Whatever option they choose, classification entails distortion of the original data. Of course, the lesson is to never make computations on the basis of map data, but to go back to the original statistics.

Choropleth maps that portray surface-related topics like densities (such as the number of inhabitants per square kilometre) do not distort the information, but choropleths of non-surface related items really present similar distortion as that described earlier for chorochromatic maps. Their false suggestion of homogeneity was first pointed out to me by Professor Eduard Imhof, when he lectured on the recently completed Swiss National Atlas in Utrecht University in 1965. He had a wonderful collection of 6 by 6 glass slides, amongst others of a population density map on which he pointed out the Luzern municipality at Lake Lucerne. Part of this municipality is situated on the other bank of Lake Lucerne, called 'am Bürgenstock' (see Figure 25.7). Of course it would have the same dark density tint as Lucerne proper, of which it is part, denoting a high population density, but, actually, no one lives in this exclave. This is demonstrated by the population distribution map in the same national atlas.

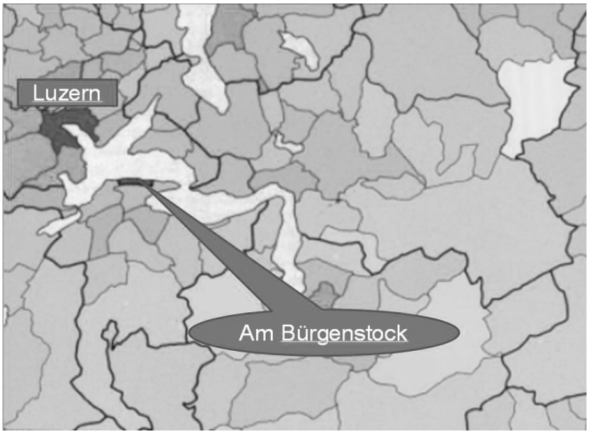


Fig. 25.7 Extract from the population density map in the Swiss National Atlas (Imhof 1967-1982)

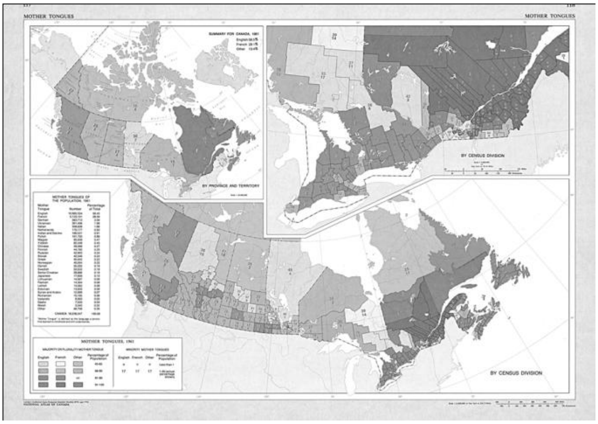


Fig. 25.8 Mother tongues in Canada (1981), from advertisement for the 5th edition of National Atlas of Canada

To underline this false suggestion of homogeneity I also show here a map of the percentage of French-speakers in Québec – as one can see (Figure 25.8), this percentage reaches over 80%, as indicated by the blue tint. This suggests the distorted information, that the whole of Québec is full of French-speaking inhabitants, while in fact the northern part of that Canadian province is almost uninhabited, except for a few Algonquin-language-speaking Indians. One thus sees how the relative choropleth maps actually need to be corrected, either by delimiting the boundaries of the ecumene, the inhabited world, or with maps that show absolute quantities - but that rarely happens, perhaps because people are not aware of that kind of information distortion.

Choropleth maps form one of the most used map types in cartography, but they do not provide a stable image – as soon as we change the aggregation level of the enumeration areas on our thematic maps, and go from postal code areas to municipalities, or districts or provinces and départements, the image of our topic can change completely. At a postal code or neighbourhood level, local differences in relative value are often so big that they will obliterate any regional or national trends. These will only become visible at higher aggregation levels, preferably for enumeration units where urban and rural municipalities have been combined. But we must also realise that an effect of aggregation is that all relative values will be smoothened.

Fig. 25.9 Influence of changes in aggregation level

At a postal code or neighbourhood level, political parties may well end up with 50-70 per cent of the votes, but this is unthinkable at higher aggregation levels, where larger absolute numbers are involved. You may ask yourself whether this is still intentional distortion. Well, because of the fact that there is only limited space for the map, and that it is consequently only possible to render one portrayal of the data, cartographers are forced to opt for one of these aggregation levels and discard the others, even if their portrayals are also valid: see Figure 25.9.

Another aspect of the choropleth map no one ever tells people about, is that it (dis)favours marginal, agricultural areas, with their skewed population composition due to the fact that young people have left the areas in order to look for jobs elsewhere, leaving a disproportionate number of elderly people. In such areas the class values will always be most extreme, for practically all socio-economic phenomena (as in unemployment maps), although the numbers of people concerned are relatively small. So the visual influence of the darker shadings for these areas is rather disproportional (see Figure 25.10, right). A similar bias is inherent in proportional point symbol maps, which always seem to favour urban areas. It seems logical that it is in these urban areas that one will find the largest symbols. Now, if I were minister of employment or labour, but had never had any training in map use, and had the relative map of unemployment in the Netherlands at my disposal, where would I invest money in order to alleviate the situation? According to the choropleth map, the highest unemployment values (of course relative values) are to be found in the north, east and south of the country, so that is where I would initiate job creation programmes. On the other hand, if I had the proportional point symbol map of absolute unemployment, based on the same data, I would instead invest my money in the Randstad, the western Holland conurbation. As with Marshall McLuhan's famous slogan "The medium is the message", here the map type we opt for is the message. We know in advance that the choropleth map will show the highest values for marginal areas, but we still produce this choropleth map because it is easy to do so, and delivers an image that is easy to interpret, but we do not warn our customers sufficiently.

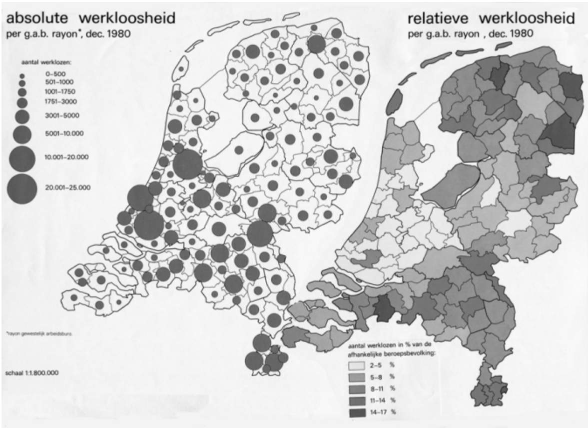


Fig. 25.10 Absolute and relative unemployment in the Netherlands in 1980 (Ormeling en Van Elzakker 1981)

4.3 Proportional symbol maps

In the 1960s and 70s cartographers first started to do perceptual research on maps – when they were about to apply the results, they were distracted by the onset of automation, and it is only in the last 10 years that perceptual studies are back on the cartographic research agenda, especially in the guise of usability studies. Even if most of the early perception studies have been forgotten, however, they gave us an insight into the distortive properties of proportional symbol maps. After psychophysical testing had been introduced (a test procedure in which the psychological reaction to physical stimuli was assessed) in order to find out whether absolute values could be better represented by proportional circles or by bars, in the famous scientific "battle of the bars and circles" that waged in the 1950s among American perception or cognitive psychologists, as a side effect it was discovered that we are unable to correctly assess the relative differences between differently sized circles. The perception psychologist Stevens established 'Stevens' Power law' according to which, if you wanted one circle to appear twice as large as another, it had to be drawn more than twice as large (Stevens 1957). If you wanted circles, bars, squares and triangles to appear equally sized, their real size had to adjusted as well. This fits beautifully with my concept of cartography as intentional distortion: instead of using proportional symbols to visualise different values this very proportionality had to be distorted in order to gain the desired effects. We embarked on bouts of psychophysical testing, following enthusiastically in the footsteps of Antonin Kolacny in Prague (Kolačny 1969), establishing minimum dimensions for specific symbols to be recognised as such until we found out that Stevens Law, on which these endeavours were based, did not apply to everyone : the exponent with which the size of symbols had to be increased in order to have the desired perceptual effects was discovered to be an average value, and the standard deviations from this average were so high that the percentage of map users who would actually benefit from the proposed corrective distortions was not particularly high. This line of research was thus discontinued.

4.4 Cartograms

In order to prevent the distortions that are caused by the portrayal of non-surface related phenomena on chorochromatic and choropleth maps, we have often taken refuge in cartograms, anamorphoses or APT (area proportional to) maps. As the enumeration areas are drawn proportional to the size of the subject's value or number for each area (and not to their geographical surface area), the resulting new base map no longer causes distortion in gauging the correct numbers. We must take care not to throw out the baby with the bath water, however, as these cartograms have as a corollary that many map users are not able to recognise the areas they represent. If we want to retain the outer limits of our country, than there will be a distortion beyond recognition of the inner regions, and if we retain a more coherent internal structure, recognising the resulting monstrosities that represent, for instance, Canada or Greece (Figure 25.11), is frequently beyond the imagination of even more advanced map users.

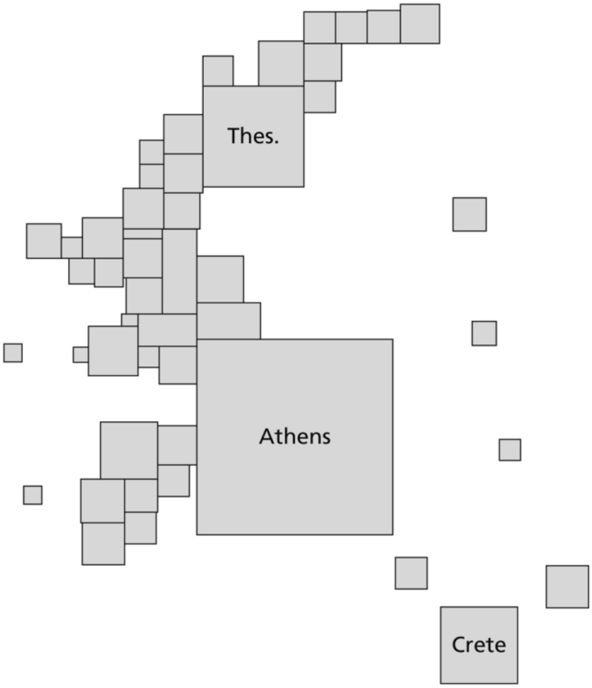


Fig. 25.11 Cartogram of Greece based on population numbers (Ormeling 2004)

5 Map Design

If mapping methods can distort reality, so can graphic design, by which term layout decisions in particular are meant here. We distort maps in our atlases because we adhere to conventions that dictate that we should use layer zones instead of landcover tints – it is only the atlases of the publishers Klett and Esselte that use these landscape tints. Children using other school atlases later have a rude awakening when they arrive at countries coloured a lush green because they were under 200m height, but actually are some of the worst deserts in the world, such as the Emirates in Arabia or the Danakil depression in Ethiopia.

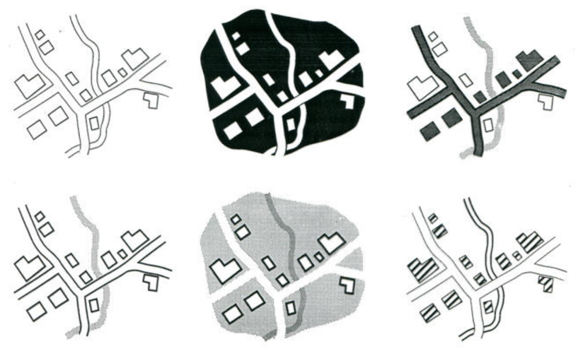


Fig. 25.12 Changing graphical emphasis

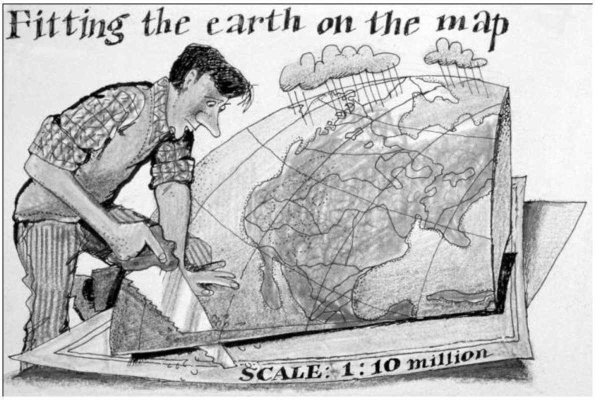


Fig. 25.13 Fitting the Earth on the map (A.Lurvink)

Map design also entails changing the emphasis on specific data categories in order to make visible their relationships – like lords of the universe we thus intervene in the landscape, we do away with whole data categories at will, in order to make some point. We may highlight the relationships between roads and buildings when studying accessibility, or between rivers and buildings when we want to assess danger of flooding (Figure 25.12). The victims that have not been selected are data categories thrown-away, that are probably roaming somewhere in cyberspace.

6 Map Projections

Our most blatant distortion of reality is of course caused by map projections – if anywhere cartographers distort reality, it is here (Figure 25.13). This is done, of course with the best of intentions, as in nearly all the cases I have described, for instance for acquiring more space on the map for rendering a specific topic, such as the flight transport of goods by the national carrier in the Netherlands. In order to visualise the goods transport which mainly occurred in the middle latitudes of the northern hemisphere, that zone had to be increased in scale in relation to the other latitudes, otherwise the symbols would not fit and also be legible on the resulting map.

Normally every projection aims to retain specific characteristic qualities of the globe, such as conformity, equivalency and equal distances, but unfortunately on a flat piece of paper they seem to rule each other out. Some projections that retain one specific characteristic do so at the expense of enormous distortions regarding the other qualities. The Peters projection (Peters 1983) has retained equivalency, but at the cost of a great loss in conformity. Other projections are able to retain equivalency without this extreme conformal sacrifice. We are also seldom warned about the reasons for the distortion in map projections, and about its benefits – although the latter may seem obvious, as in the case of the equal area representation of the world's oceans by National Geographic (National Geographic 2012), the drawbacks of such distortions are not mentioned.

7 Distortion by Abstraction

Finally, we distort because the Earth's surface is abstracted, because we make invisible our principal subject, humankind, usually in order simply to prevent clutter on our maps. Sometimes, on old town plans, people are still visible but mostly the reduced scale makes it impossible to see them. Nor do we see cars or trucks on our roads, trains on our railroad tracks, aeroplanes or ships, not even on our animated maps. This in my opinion should be addressed by new generations of cartographers as they programme new digital maps and atlases.

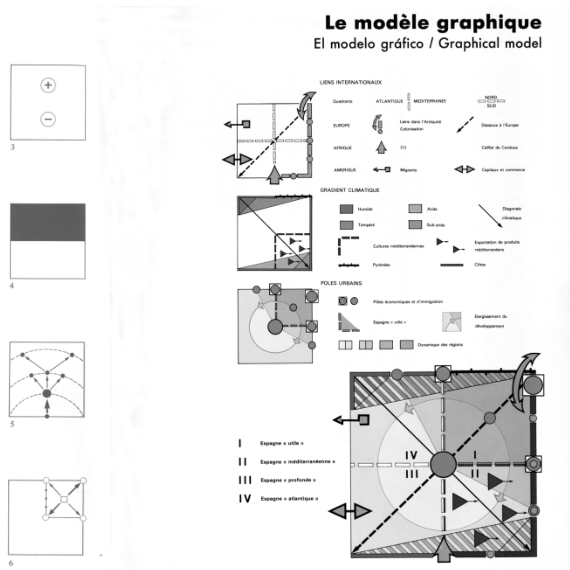


Fig. 25.14 From maps to chorèmes: abstraction of the map image (Ferras 1986)

At least the intentions of the human agent are accounted for in chorèmes, scheme-like maps with their own symbology that expresses human forces such as expansion and retreat, attraction and repulsion, fronts and hierarchies. Although geographical space is reduced or distorted here as much as it could be, to geometrical shapes such as squares and hexagons, at least one is no longer distracted by random geographical objects and one is able to grasp the essential influences, forces and trends that are at work in our societies. For instance, the French economy can be explained by different demographical dynamics and attitudes towards investment in manufacturing, by highly dynamic innovative industries on the Mediterranean and Alpine boundaries, and by an outdated industrial zone in the northeast that needs to be restructured. All these forces, expressed as so many chorèmes, can be combined in order to present an overall scheme of the French economy. The same abstraction of human forces and geographic space can be applied to any other society, like Spain for instance, and have a beneficial effect on our understanding, even if at the same time the link with reality is threatened: it is good if we understand what makes the Spanish economy tick, but if we simultaneously lose the facility to locate these productive forces then the whole exercise might be pointless (see Figure 25.14).

8 How Can We Fight the Effect of These Distortions?

Now how can we profit from all these observations on distortions? Have these different kinds of cartographic distortions, all well-meant but all detrimental as well, something more in common? Their major common characteristic is that there is always a trade-off. When cartographers visualise some specific characteristics, links, relationships or notions, then others must be left out. If the maps have been made relevant to specific user groups, they lose their relevance for others. Serving specific purposes, such as better relief representation or achieving individual or overall clarity for some phenomena can only be reached if we leave out many other spatial data categories. When we opt for specific map projection characteristics, severe distortion of other projection characteristics will result.

This all means – because map users are insufficiently aware of these corollaries of distortion – that there should be more warnings attached to maps regarding these drawbacks, and the unsuitability of the use of these maps for specific purposes, such as making measurements, should be clearly stated. It should be easy to attach such warnings to maps, especially to their digital files.

The other action to take is to provide corrective cartographic measures, for instance by adding proportional symbols related to the relevant absolute quantitative data to both chorochromatic and choropleth maps of non-area related data. In general: we should warn about the costs of the distortions we cause, and try to offset their unwelcome consequences. Even if we cartographers know how we have distorted reality, the users of our maps may not know it. Democratization of mapping due to the availability of software mapping packages now allows everyone to produce maps and to follow cartographic practices, but the inherent knowledge of the cartographers of the side effects of their distortions has not been democratised at the same time. Our input remains necessary to set things right!

References

Ferras R (1986) España-Espagne-Spain. Atlas RECLUS, Paris: Fayard/Reclus.

Garland K (1994) Mr Beck's Underground Map. London: Capital Transport Publishing

Imhof E (Ed.) (1961-1982) Atlas der Schweiz. Wabern: Eidgenösssische Landestopographie.

Jenks GF (1967) The data model concept in statistical mapping. International Yearbook of Cartography 7: 186–190

Klee P (1920) Schöpferische Konfession. Berlin: Erich Reiß Verlag.

Kolačny A (1969) Utilitarian cartography – the road towards the optimal effect of cartographic information. ICA Working Group on Cartographic Information, Prague.

Knöpfli R (1992) Die natürliche und die abstrakte Landschaft. Symposium Wissenschaft und Technik im künftigen Deutschland.

Krallert W (1961) Methodische Probleme der Völker- und Sprachenkarten. Internationales Jahrbuch für Kartographie Vol.1 pp 99-121. Gütersloh: Bertelsmann Verlag.

Kranz F (1966) Die Generalisierung für eine achtfache Verkleinerung. pp 99-119, in H.Bosse (Ed.) Kartographische Generalisierung. Niederdollendorf 1966. Mannheim: Bibliographisches Institut.

National Geographic (2012) “Beneath the Oceans” map supplement, September 2012 issue, Washington: National Geographic

Ormeling F (2004) Island Cartography. In: Nisiotiki Chartografia. Proceedings 7° Ethniko Synedrio Chartografias, Mitilini 2002. Pp 24-30. Thessaloniki 2004.

Ormeling F, Van Elzakker C (1981) Geregistreerde werkloosheid in Nederland 1980/81. Kartografisch Tijdschrift 7, 3: 53-56.

Ormeling FJ (1964) De Grote Bosatlas. Groningen: Wolters-Noordhoff.

Peters A (1983) Die Neue Kartographie/The New Cartography Klagenfurt, Austria: Carinthia University; New York: Friendship Press.

Stevens SS (1957) On the psychophysical law. Psychological Review 64 (3): 153–181

Westermann Verlag (2010) Diercke Weltatlas, Braunschweig: Westermann.