Chapter 11
Generalisation in Practice Within National Mapping Agencies

Cécile Duchêne, Blanca Baella, Cynthia A. Brewer, Dirk Burghardt, Barbara P. Buttenfield, Julien Gaffuri, Dominik Käuferle, François Lecordix, Emmanuel Maugeais, Ron Nijhuis, Maria Pla, Marc Post, Nicolas Regnauld, Lawrence V. Stanislawski, Jantien Stoter, Katalin Tóth, Sabine Urbanke, Vincent van Altena and Antje Wiedemann

C. Duchêne (✉)
Laboratoire COGIT, IGN, Saint-Mandé, France
e-mail: cecile.duchene@ign.fr

B. Baella · M. Pla
Institut Cartogràfic de Catalunya, Barcelona, Catalonia, Spain
e-mail: blanca.baella@icc.cat

M. Pla
e-mail: maria.pla@icc.cat

C. A. Brewer
Department of Geography, Pennsylvania State University, University Park, Pennsylvania, USA
e-mail: cbrewer@psu.edu

D. Burghardt
Dresden University of Technology, Dresden, Germany
e-mail: dirk.burghardt@tu-dresden.de

B. P. Buttenfield
Department of Geography, University of Colorado, Boulder, Colorado, USA
e-mail: babs@colorado.edu

J. Gaffuri · K. Tóth
Institute for Environment and Sustainability, Joint Research Centre, European Commission, Ispra, Italy
e-mail: julien.gaffuri@gmail.com

D. Käuferle
Federal Office of Topography swisstopo, Wabern, Switzerland
e-mail: dominik.kaeuferle@swisstopo.ch

D. Burghardt et al. (eds.), *Abstracting Geographic Information in a Data Rich World*, Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-00203-3_11,

Abstract National Mapping Agencies (NMAs) are still among the main end users of research into automated generalisation, which is transferred into their production lines via various means. This chapter includes contributions from seven NMAs, illustrating how automated generalisation is used in practice within their partly or fully automated databases and maps production lines, what results are currently being obtained and what further developments are on-going or planned. A contribution by the European Joint Research Center reports on the use of multiple representation and generalisation in the context of the implementation of the European INSPIRE directive. The chapter finishes with a synthesis of recent achievements, as well as future challenges that NMAs have begun to tackle.

F. Lecordix · E. Maugeais
Developments department, IGN, Saint-Mandé, France
e-mail: francois.lecordix@ign.fr

E. Maugeais
e-mail: emmanuel.maugeais@ign.fr

R. Nijhuis · M. Post · J. Stoter · V. van Altena
Kadaster Geo informatie, Zwolle, Netherlands
e-mail: ron.nijhuis@kadaster.nl

M. Post
e-mail: marc.post@kadaster.nl

J. Stoter
e-mail: jantien.stoter@kadaster.nl

V. van Altena
e-mail: vincent.altenavan@kadaster.nl

N. Regnauld
Ordnance Survey Great Britain, Southampton, UK
e-mail: nicolas.regnauld@1spatial.com

L. V. Stanislawski
Center of Excellence for Geospatial Information Science,
United States Geological Survey, Rolla, Missouri, USA
e-mail: lstan@usgs.gov

S. Urbanke · A. Wiedemann
AdV Project, LGL, Stuttgart, Baden-Württemberg, Germany
e-mail: sabine.urbanke@lgl.bwl.de

A. Wiedemann
e-mail: antje.wiedemann@lgl.bwl.de

11.1 Introduction

11.1.1 Generalisation, from Research to Production

Research in generalisation has long been driven by the needs of National Mapping Agencies (NMAs).[1] NMAs remain the primary users of automated generalisation processes, although some generalisation functionalities are currently made available to more and more geographic information users via GIS software and the web.

Various schemes exist within NMAs to encourage research in this area, and then to transfer research results into production environments. Some NMAs include long-term research teams to which they can directly express future needs (e.g. IGN-France, USGS-USA, OSGB-Great Britain). This research is sometimes undertaken in collaboration with universities. Other NMAs do not have long-term research teams but collaborate with universities while subcontracting long term research tasks related to their needs (e.g. Kadaster-Netherlands, ICC-Catalonia). When it comes to industrialisation of research results, some NMAs have development teams that develop generalisation tools specific to their needs, partially based on previously published research results, usually on top of commercial off the shelf GIS software (e.g. ICC-Catalonia, Kadaster-NL, IGN-France, OSGB-Great Britain). Other NMAs subcontract the customisation of existing GIS software to GIS vendors, who also have "research and development" teams (e.g. Swisstopo-Switzerland, AdV-Germany).

In addition to this, events are organised where researchers, practitioners from NMAs and GIS vendors can meet and exchange ideas and information. The International Cartographic Conference, every two years, and the annual workshop of the ICA Commission on Generalisation and Multiple Representation, are good examples of such events. This enables NMAs to articulate their needs, researchers and vendors to understand needs, and vendors and developers from NMAs to keep up to date with recent research achievements. Other meetings have also been organised specifically for NMAs, so that they can share experiences and identify common needs that can then be communicated to the research and vendor communities (e.g. Stoter 2005; ICA and EuroSDR 2013). The survey by Foerster et al. (2010) also identified shared needs among NMAs.

Collaboration between researchers, NMAs and vendors has also been undertaken as projects. Some projects seek to build advanced prototypes on top of existing commercial or homemade software. Examples include European projects such as AGENT (Barrault et al. 2001), GiMoDig (Sarjakoski et al. 2005), ESDIN (Kereneur et al. 2011), and ELF (Jakobsson et al. 2013). A second kind of project aims at evaluating existing generalisation solutions through benchmarking: the OEEPE test (Ruas 2001), and the EuroSDR project (Stoter et al. 2009b, 2010).

[1] This acronym is a misnomer since some of them are national while others, within federal countries, can be regional. However, it is widely used in the ICA Community, therefore our choice to use it in this chapter.

Such tests not only enable evaluation of existing solutions, but also result in improvements, both in the short term (because the software providers were allowed to adjust the software within certain rules in order to complete benchmark tasks), and the long term (identifying shortfalls helps define future research agendas).

Finally, a last opportunity for NMAS to see their needs taken into account are the competitive tender processes they run when they renew their production lines. Indeed, during such tender processes, vendors are given the opportunity to tailor their software according to the requirements of the concerned NMA, and they often take this opportunity in order to maximise their chances of retaining contracts.

Thus NMAs gain new operational generalisation functionalities that meet their needs, both through regular improvements of generalisation functionalities of commercial GIS software, and through ad hoc customisation that can be made either from within the organisation, or by subcontracting to GIS vendors. A large part of those improvements and customisations rely on published research such as the ones presented in various chapters of this book.

11.1.2 Outline of the Chapter

The next seven sections of this chapter describe contributions from seven of the most active NMAs within the ICA Commission in Generalisation and Multiple Representation. Together they represent a broad selection of current users of automated generalisation in map production. They are ICC, Catalonia (11.2), IGN, France (11.3), Swisstopo, Switzerland (11.4), OSGB, Great Britain (11.5), USGS, USA (11.6), Kadaster, Netherlands (11.7), and AdV, a consortium group of several regional mapping agencies in Germany (11.8). Each section gives insights into the current use of generalisation and multi-representation (the main strategy of use and technical details for specific production lines), results achieved so far and future plans of the NMA. Section 11.9 reports on pieces of work conducted at the European Joint Research Center (JRC) to support the implementation of the INSPIRE European directive in European NMAs, regarding multi-scale modelling on the one hand, and the setup of generalisation services on the other. Finally, Sect. 11.10 provides a synthesis of recent achievements, current trends, on-going work and future challenges, based both on the contributions from NMAs included in this chapter and on information gathered by direct interactions with NMAs during a symposium held in March 2013 (ICA and EuroSDR 2013).

11.1.2.1 Notice to Reader

Please be aware that the facts reported in this chapter are up to date in 2013, but might evolve quite quickly since the developments in generalisation are currently particularly active in several NMAs, as a consequence of the changing context in the geoinformation world in general (see Sect. 11.10.2 and Chap. 12 of this book), and the maturity reached by research in generalisation linked to the traditional needs of NMAs (topographic data and maps production). In addition, please note that some concepts and ideas referred to by the contributions from NMAs (Sects. 11.2–11.8), such as the distinction between DLM and DCM or the star vs. ladder derivation schemes, are further elaborated in Sect. 11.10.2.

11.2 Deriving Products Through Generalisation at the Institut Cartogràfic de Catalunya

Maria Pla and Blanca Baella

Since the foundation of the Institut Cartogràfic de Catalunya (ICC) in 1982, one of the main activities was the production of georeferenced data, mainly topographic and thematic data and orthophoto images. The derivation of products at smaller scales using generalisation techniques started in 1996, after the first digital coverage at a scale of 1:5,000 of Catalonia was completed. Although initially the generalisation techniques were applied only for the derivation of vector topographic maps, some years later they were also implemented in the derivation of vector databases, raster orthoimages and map names. Generalisation workflows of topographic data and map names are semiautomatic processes that use commercial software enriched with software developed at the ICC. In the case of orthoimages the workflow is completely automatic using applications developed at the ICC.

In the last few years, high demand for derived products for visualisation via the internet and over mobile devices has introduced new requirements, both in terms of the data and in the processes used, demanding more intelligent data models and on-the-fly generalisation. The current ICC on-going work is mainly focused on these two topics.

11.2.1 Generalisation of Topographic Data

11.2.1.1 Current Production

Topographic databases at different resolutions are produced and maintained ranging from 1:1,000 scale to 1:25,000 scale. The most detailed information covers the urban areas at 1:1,000 (20 cm of accuracy), while data at 1:5,000 (1 m accuracy) and 1:25,000 (2.5 m accuracy) covers the whole Catalonia. These databases are compiled in 2.5D using stereo plotting on top of digital photogrammetric systems.

Smaller scales such as 1:50,000 and 1:250,000, also covering the whole territory, are collected by digitizing in 2D on top of orthophoto images.

After a first generation of spaghetti data collected using CAD systems, more complex models were designed and implemented that enable new data exploitation using GIS systems. The topographic data models were designed to preserve the semantic coherence between scales, but without explicit relationships between the representations of the same geographical object in the different databases. This was because of the lack of commercial tools to maintain them and because of the different update cycles of each product. The generalisation techniques were applied to the Topographic Database at 1:5,000 (BT-5M) to derive the Topographic Map at 1:10,000 (MT-10M) and the Topographic Database at 1:25,000 (BT-25M), and to the Topographic Map at 1:50,000 (MT-50M) to derive the Topographic Map at 1:100,000 (MT-100M).

The BT-5M is compiled using photogrammetric systems according to a 2.5D data model and stored in DGN files from MicroStation. During the stereo plotting process all the features required to generate a digital terrain model (DTM) and a digital surface model (DSM) are compiled together with the topographic objects. The updating cycle is 5 years over the entire territory and more frequently over the most dynamic areas, located mainly near the coast. The first compilation of the BT-5M started in 1985 and was completed in 1995. At that time, the underlying data model was based on 2.5D “spaghetti” vectors. It supported the generation of DTMs, but it was not designed for GIS. The next version of the model addressed this shortcoming. Moreover, it supported automatic generalisation and 2.5D topographic objects.

The MT-10M is just a map, not a database. It contains 2D data obtained by semiautomatic generalisation from the BT-5M. Automatic processes include elimination of some objects, class aggregation, building simplification, altimetric point selection and selection and scaling of map names. Manual editing is applied to refine the automatic results, but because the scale change between the original and the target scale is quite small, it can be performed in few hours, around 20 h per sheet (3,200 ha). The updating cycle is the same as for the original database and the updating workflow generalises again the updated BT-5M without keeping any object of the old MT-10M, because most of the generalisation operations are automatic and the manual editing is quite fast.

The need for an updated base data at 1:25,000, the existence of a production program for the BT-5M, and the experiences at the ICC in implementing generalisation workflows, allowed us to start producing the first version of the BT-25M using generalisation processes in 2003 (Baella and Pla 2003). Compared with previous ICC generalisation experiences, the workflow entailed two challenges: to obtain a topographic database, not only a map, and to generalise 2.5D data instead of 2D data. The main difference between generalisation for obtaining a map or for creating a database comes from having to preserve the topological structure of the

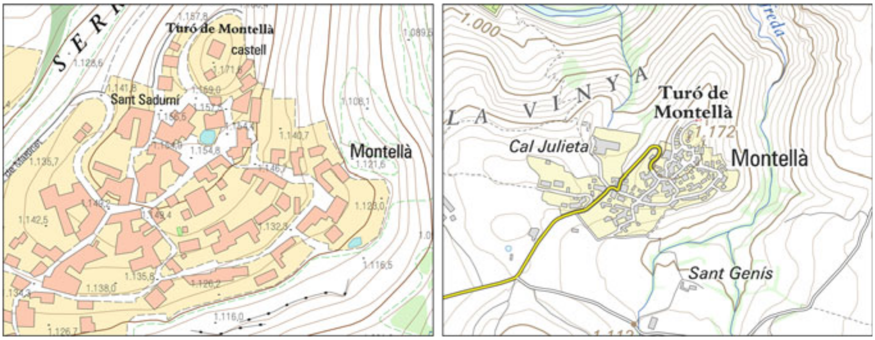

Fig. 11.1 At the *left* the original BT-5M data, at the *right* the resulting BT-25M after generalisation

data and their attributes. The 2.5D characteristic of the generalised data required new software development and a careful editing process (Baella et al. 2007; Palomar-Vázquez et al. 2011).

The BT-25M is obtained by generalisation of the BT-5M (Fig. 11.1), is updated using stereoplotting of recent photogrammetric flights and it is completed with data extracted from thematic databases and, occasionally, with data collected in the field. Automatic generalisation includes the same operations used for the MT-10M, plus the simplification of linear elements and the generalisation of the elevation of the objects. Manual editing is used for generalisation operations that cannot be automated such as typification, exaggeration, collapse, aggregation or refinement of the automatic results. This requires an average of 150 h per sheet (12,800 ha). The database is updated independently of the original database, because the updating cycle planned for the BT-25M is approximately 2 years, much shorter than the updating cycle of the BT-5M.

The MT-100M is obtained by applying 2D generalisation to the MT-50M. Automatic processes include selection, class aggregation, line simplification and selection and scaling of map names. Manual editing is applied to generalise buildings and to refine the results.

The generalisation workflows described above (summarised in Fig. 11.2) combine commercial software and ICC developments. For some specific aspects, the ICC collaborates with external research groups. The commercial software CHANGE, developed and distributed by the University of Hannover, is used for the building generalisation up to 1:25,000 scale. The ICC developments are mainly focused on line and map names generalisation, and in the tools for supporting interactive generalisation, for example collapse, exaggeration or conflict detection and resolution. Tools for terrain generalisation, as spot height selection, have been developed in collaboration with the Department of Cartographic Engineering, Geodesy and Photogrammetry of the Universitat Politècnica de València.

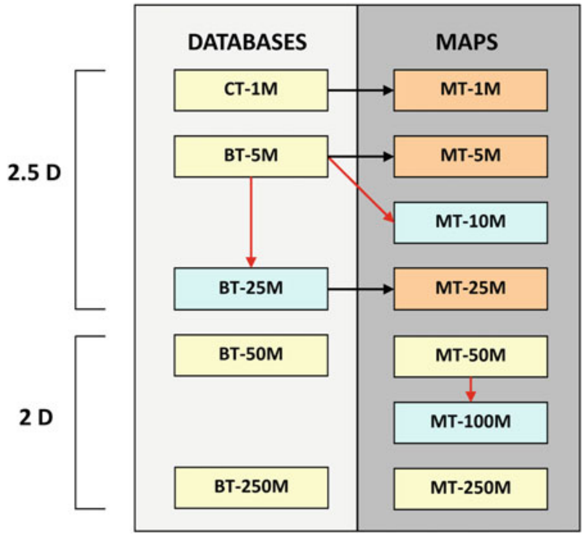

Fig. 11.2 General schema representing the existing ICC topographic databases and maps. Products compiled from primary sources are indicated in *yellow* and products derived applying generalisation in *blue*. Products in orange are derived by automatic symbolisation without any editing process

11.2.2 Generalisation of Orthophotos

The ICC has a long tradition in orthophoto production using software developed internally. Since its foundation a large collection of orthophoto products at different scales has been produced using aerial or satellite images, ranging from 10 cm to 25 m pixel ground size. At the beginning, in the 1980s, each orthophoto product used its own original image source, but in the year 2000, the production of the orthophoto of Catalonia at 50 cm pixel ground size was changed in order to ensure not only the geometrical continuity but also the radiometric continuity over the entire territory. This offered the possibility of using generalisation processes to derive orthophoto products with a larger pixel ground size.

The generalisation processes applied on the orthophotos implies, as in the case of the topographic data, a scale change, which is implemented through a reduction in the number of pixels of the image and through the change of the ground resolution size of the pixel. Basically this process has two key aspects: the determination of the new pixel position, which is a geometric operation, and the calculation of the radiometry of the new product, applying a convolution using a Gaussian bi-dimensional function for eliminating the higher frequencies of the radiometric values of the resulting image and ensuring a good radiometric result.

These generalisation processes applied to the orthophotos have optimised the production of image products at the ICC allowing a higher rate of productivity. Nowadays, from the base product, the orthophoto of Catalonia of 25 cm pixel ground size, the ICC is deriving the orthophotos of 50 cm and 2.5 m (Fig. 11.3) covering all Catalonia on an annual basis.

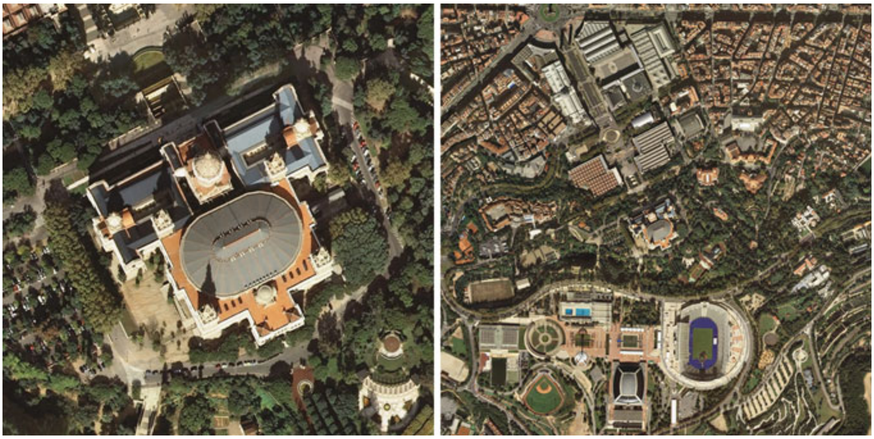

Fig. 11.3 At the *left* the original orthophoto at 50 cm data, at the *right* the resulting at 2.5 m after generalisation

11.2.3 Generalisation of Map Names

From its foundation, the ICC has compiled and maintained a very detailed collection of map names. The first data compilation was made through a field campaign, using orthophotos as a base reference document, started in the year 1984 and finished in 1991. The information was compiled and georeferenced on top of the orthophotos manually in the field. Following this, the alphanumerical information associated with the map names was translated to a digital text file. The collection of map names was then used in the publication of several ICC products, for example the Topographic Map of Catalonia at 1:5,000 or the orthophotomap at the same scale. Since this first campaign, the ICC has a continuous updating process using all the sources available, from new smaller field campaigns to external data incorporation. The current data model of the ICC Map Names collection is based on a set of graphical data and a set of alphanumerical information linked through an identifier. But the map names are not linked with the elements stored in the topographical database. The new topographic data model designed for MRDB purposes will include this link.

Due to the widespread use of map names in cartographic products, from the beginning there was a need for map names generalisation tools. These tools were developed at the ICC, at the beginning of the 1990s, and since then they have been continuously improved (Baella and Pla 2001).

Two types of generalisation tools were developed for map names: the automatic tools and the interactive ones devoted to helping the cartographers in the improvement and refinement of the automated results. The automatic tools include the map name selection and the new graphical placement of the map names, following the style and the scale of the new product. The interactive tools include (1) conflict detection of the generalised name placement with other map name placements or with topographical elements, (2) the detection of placements of the same map name too close together, and (3) placement style improvement, for example the space between different text lines or the base supporting text lines.

The customised generalisation tools developed at the ICC have enabled us to derive collections of map names while optimising production costs and ensuring coherence between the original and the derived collection.

11.2.4 On-Going Work

The huge demand for updated information and the high pressure to obtain derived products for visualisation via the internet and mobile devices has introduced new requirements, especially related to topographic data, that are costly to achieve with the current data models. At this moment the ICC is assessing a solution based on the implementation of a MRDB for integrating topographic data at different scales that allows to optimise generation of derived products using generalisation, including on-the-fly processes for intermediate resolutions used mainly for visualisation. The implementation should take into account the current state of the ICC databases, the existing limitations in commercial software, the requirements to be implemented in the ICC production environment and the challenge of achieving reasonable productivity ratios.

Although the MT-10M and the BT-25M are derived by generalisation from the BT-5M, there are no explicit mechanisms linking the original and the generalised datasets. The main reason for this is that when the workflows were implemented, there was a lack of tools for managing the links during the compiling photogrammetric process. In the last few years, GIS based photogrammetric tools ready for production environments and delivering reasonable productivities have started to become common. The availability of these tools has encouraged the ICC to design a new version of the data model that would implement the main aspects of a multirepresentation database (MRDB), linked objects at different levels of detail, where the links between the objects will be established through unique identifiers (Baella et al. 2012). In the new data model, each feature instance will be characterised by an identifier that must be unique and persistent throughout the whole life-cycle of the feature instance and never reused.

The ICC MRDB will be based upon one single schema with linked features, where one feature belongs to one single resolution and has a link with one, or more, features at the other scales. The links between corresponding feature instances at different levels of detail will be determined by the various cardinalities that exist in the relations between them. Due to the complete coverage of the current version of the BT-5M and the BT-25M, the links between the feature instances at different resolutions will be established using matching techniques that will be applied in the migration of the existing data to the MRDB data model. For later updating processes on the higher accuracy data, the lower resolution data and the links will be managed during the generalisation process.

The implementation of the new ICC topographic data model will involve two main tasks, the migration of the photogrammetric data collection from a CAD system to a GIS system based on a DBMS, and the management of the data and the MRDB relationships. The migration to a GIS environment will require customisation of the commercial system and the training of the production teams in the GIS system and in the customised tools. The customisation includes the development of a set of tools that preserves the 2.5D nature of the data model and achieves a similar level of productivity to that of the existing CAD environment. The management of the MRDB links will be based on tables that store the relationships between features instances through unique identifiers. The first implementation of the MRDB will be done from the two existing databases, BT-5M and BT-25M, which do not have any explicit link between the feature instances representing the same geographical object. The matching processes to find the related feature instances will fill the tables of relationships. In further updating processes of the database, the generalisation operations to derive the smaller scale must automatically establish and manage these links.

11.3 The New Base Map Project: A Semi-Automated Production Line for Topographic Maps at IGN-France

François Lecordix and Emmanuel Maugeais

11.3.1 Important Investment in R&D at IGN-France

At the beginning of the 1990s, IGN France decided to launch a large and long term investment into research and development in generalisation. This investment was justified by the need to solve the problems of producing maps at different scales directly from databases that IGN had already produced and updated for a number of years. The main goal for the French national mapping agency was to reduce costs and, in particular, only collect the data once and use them to produce all other databases and maps. This aim stipulated that maps have to be derived from databases and the cost of deriving them should be affordable which means the process should be as automatic as possible.

In 1991, IGN placed the COGIT laboratory in charge of research on generalisation with the aim of deriving DCMs from DLMs. A team of four researchers was initially nominated and there have been between four and eight researchers working until 2013 on this project. During this period, 13 PhDs have defended their theses and many results were obtained on different aspects of generalisation:

platforms and algorithms (e.g. Lecordix et al. 1997), data enrichment (e.g. Boffet and Coquerel 2000), strategies (e.g. Ruas 1999), evaluation (e.g. Bard 2004). A network of strong collaboration was established with different universities and NMAs around the world. A part of this research was capitalised during the European project AGENT [1997–2000, see (Barrault et al. 2001)] in a generalisation prototype built on the commercial software LAMPS2 (from Laser-Scan company, now called 1Spatial, who took part in the project).

During the AGENT project, in June 1999, IGN launched the "Carto2001" development project with 4 developers, with the aim of setting up a new production flowline to derive a 1:100,000 DCM from a 10 m resolution DLM, BDCarto®. In 2004, the Carto 2001 project provided automatic solutions for networks generalisation, using agent technology for road bends coalescence and the "Beams" algorithm (Bader 2001) for overlapping in symbolised networks. The flowline was developed on LAMPS2 editor, which provides the Gothic DBMS, the AGENT prototype and many other developments, either by 1Spatial or homemade, for generalisation and updating (Jahard et al. 2003; Lecordix and Lemarié 2007). Only automatic label placement has been fully developed in-house at IGN in the form of a software called WinPAT.

This flowline was used to derive the first edition of DCM Top100 (2006–2008) with eight cartographers and then to update this DCM twice (Plessis 2011). The production, still in use, has provided for the first time experience of generalisation in production environments and has enabled to study the introduction of new solutions for generalisation problems at other scales.

11.3.2 New Base Map Project

11.3.2.1 Different Purposes

Since the end of the 1980s, IGN has been producing the BDTopo®, a DLM with a precision of 1 m. This DLM contained all information (networks, buildings, landscape, contour lines, point, etc.) for starting from 1993, the semi-automatic production of maps at 1:25,000, in a new version called Type 93. Some choices of symbolisation were made to minimise problems with generalisation. For example the widths of symbols are very thin, the buildings are only exaggerated not displaced, and the roads layer is drawn on top of the buildings layer so that the overlaps between roads and buildings are less visible. A problem with this flowline is that maps at 1:50,000 scale (used by the military) was no longer produced because of generalisation problems. By 2000, IGN had only produced 25 % of the French territory in BDTopo®. To speed up the construction of BDTopo®, new specifications and processes were selected to finish the DLM collection for the whole territory in 2007. This decision meant that the production of Type 93 would not be possible any more: the new version of the DLM would not contain all needed information. So, in 2004, IGN launched the New Base Map Project. The main challenges were:

- defining processes to collect additional information necessary for cartographic production,
- deriving 1:25,000 and 1:50,000 seamless DCMs, incorporating human operators in an user friendly environment,
- proposing an automatic process to propagate the updates collected in DLM to DCM,
- and, last but not least, the requirement that this should be slightly less expensive than previous flowlines even if there is less information in the new version of BDTopo®.

The staff of the New Base Map Project varied between three and five developers during the period 2004–2011. The first result was obtained in 2009 and it allowed the production of the collection of additional information to start. Thus, cartographic production was ready in 2010, hence the new version of the map is called Type 2010 (Maugeais et al. 2011). During that period, some modifications were introduced in IGN vector databases. Different DLMs that IGN managed were merged into one DLM called BDUni. Some modifications introduced in BDUni had a strong impact on the cartographic process under development. For example, in the new specifications for roads, lanes of dual carriageways and motorways were recorded separately (and not merged as before in BDTopo®). For buildings, the outline of a set of buildings were gradually replaced by cadastral buildings with more details and more partitions. BDUni is managed in DBMS PostgreSQL in a seamless database and the French GIS software GeoConcept (from Geoconcept company) is used to modify the data. The communication between the DBMS and GIS is done via the software GCVS that IGN had developed.

11.3.2.2 Technical Architecture and Flowline

The New Base Map Project developers decided to retain an architecture with different software, similar to BDUni's architecture, but it has an extra software specific to cartography. PostgreSQL provides the DBMS, the interactive editor is GeoConcept, with a specific layer designed for cartography, Publisher (Guislain and Lecordix 2011), and GCVS provides the link with DBMS. The software Clarity (from 1Spatial) was selected for automatic cartographic processes and generalisation by the New Base Map Project developers who developed a specific bridge between GeoConcept/Publisher and Clarity, called CLEO, which enables to export data from GeoConcept to Clarity, to launch automatically processes in Clarity and then to send data back from Clarity to GeoConcept. The IGN solution WinPAT was used for automatic label placement.

The New Base Map Project defined the flowline for the new topographic map. A new database called BDComplementaire was defined to store the data missing in BDUni but necessary to produce maps, i.e. touristic itinerary, contour lines etc. Different processes were defined to collect this information: for example, field work, information extraction from old maps or orthoimagery.

After collection, the cartographic process can start with the first step of merging BDUni and BDComplementaire and generating a new database BDRef. This BDRef is not dependent on one scale or a particular symbolisation, but some operations are applied to compute information and to introduce consistency: building and dual carriageways are merged (see below), itineraries are conflated on topographic networks, urban structures such as settlement areas or city blocks are computed between buildings and networks. The BDRef is an intermediate DLM where model generalisation has been applied from BDUni, so that DCMs at scales 1:25,000 to 1:50,000 can be derived by cartographic generalisation.

The next step consists in deriving a DCM at a defined scale, France25 at 1:25,000 or France50 at 1:50,000. The New Base Map Project worked mainly on France25, but some experiments were made on France50. In these DCMs, the symbolisation is assigned at the beginning of the process and then many automatic generalisation and label placement processes are launched. Between these automatic processes, manual cartographic editing is applied. This full process is applied on tiles of 20 by 20 km. The objects which cross many tiles are handled in each tile and a semi-automatic process is launched to manage these objects when the cartographic data are stored in the seamless database.

11.3.2.3 Main Generalisation Operations

Four main steps are necessary to ensure efficient building and network generalisation:

1. Building merging: a building is stored in the BDUni database as a plurality of objects that resulted from dividing it by cadastral plots. Buildings that are adjacent and have the same symbolisation are merged before running the generalisation process (see below). It prevents the system from assessing them too small individually, enlarging them independently from each others and as a result generating lots of overlaps.
2. Dual carriageway merging: dual carriageways, for which two lanes are captured in BDUni, are merged using a five stage process adapted from Thom (2005). Road segments that form part of a lane in dual carriageways are detected, grouped into lanes and matched by pairs, the central axis is computed by a skeleton algorithm of the central zone between each pair of lanes, and finally connections and intersections with the rest of the network are managed. To preserve bridges and tunnels and avoid inconsistent merging, this sequence is performed one by one on five categories of roads according to their nature and their position relative to the ground.

3. Network generalisation: the solution developed during the Carto2001 project was adapted for topographic objects (e.g. embankments). Firstly the process detects groups of connected objects concerned with overlapping cartographic conflicts, called "flexibility graphs" (Lecordix and Lemarié 2007). The "beams" algorithm (Bader 2001) is used to move networks to solve these conflicts. Finally a process based on the GAEL model (Gaffuri 2007) moves buildings to maintain topological consistency with the roads.
4. Building generalisation: the AGENT model proposed by Ruas (1999) and further developed during the AGENT project (Barrault et al. 2001) has proved its worth for building generalisation. This model works at two levels: a "micro" level for individual objects (in this case, buildings) and a "meso" level for groups (here "building groups", consisting of clusters of close buildings previously created with a method based on Boffet and Coquerel (2000)). Different constraints and possible actions (generalisation algorithms) to solve them are defined at those two object levels. The agent engine ensures the orchestration of the actions to solve the constraints which are defined as follows:
    - A density constraint ensures that the black to white ratio within a building group does not increase too much during scale reduction. It determines the elimination of one or more buildings based on several criteria: the building area, its distances to the neighbouring buildings and its proximity to roads surrounding the group.
    - Constraints of minimum size, granularity and shape preservation are defined at the buildings level: they are solved by exaggeration and simplification of the buildings.
    - A proximity constraint allows the movement of buildings to correct the overlap between them (due to their exaggeration at the "micro" level) or with the symbolised roads.

11.3.3 Production Launch and Results

In 2010, experimentation of the cartographic process was launched in production for the first time (an example of output is shown in Fig. 11.4). Some improvements were required and the main part of the process was validated in 2011 to start the production on rural areas. Subsequently many improvements were made such as generalising large urban areas, correcting bugs in specific cases, managing tile edge 'reconciliation', and testing at 1:50,000 scale (e.g. Fig. 11.5).

At the end of 2012, 317 tiles were collected (for additional information) over 280 h for each and the cartographic process was run at 1:25,000 on 136 tiles in 175 h. To produce the complete DCM at 1:25,000 for the French territories (overseas included), 1637 tiles will have to be produced in ten years.

Other tools are being developed in the course of the New Base Map Project but they have not been used in production yet. However they are necessary to produce

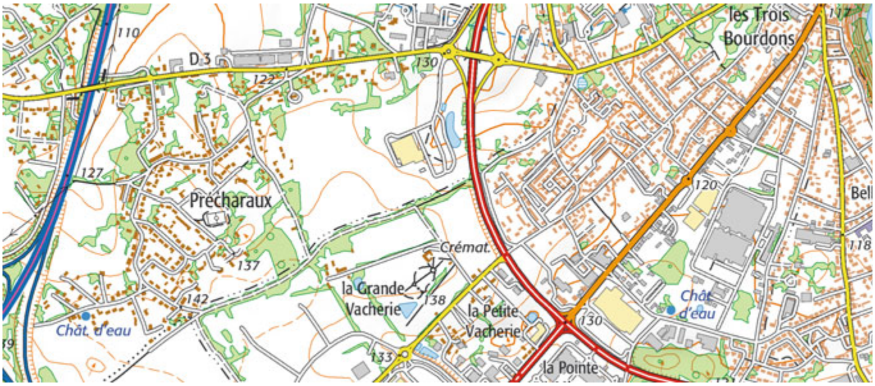

Fig. 11.4 Cut-out of base map Type 2010 at 1:25,000 scale (IGN-France)

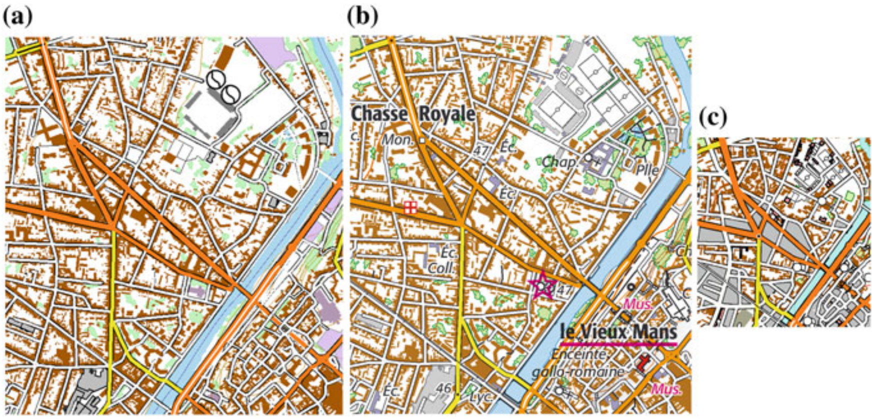

Fig. 11.5 Generalisation results of initial data (**a**) at 1:25,000 scale (cut-out of base map Type 2010) (**b**) and first trials at 1:50,000 scale (**c**)

the DCM for the whole country. These include for instance: the collapse of hedges, modelled as areas, into lines by skeletonisation (Maugeais et al. 2011), contour lines extraction from DTMs with generalisation solutions (Jaara and Lecordix 2011), and mountain representation (Gondol et al. 2008) (e.g. Fig. 11.6).

11.3.4 Conclusion and Future Works

To continue the IGN research, development and production strategy, in 2010 IGN launched the project "Map on Demand" using COGIT's research on this subject (e.g. Christophe 2011) and exploiting the results of the New Base Map project. After adaptating Type 2010 flowline, the "Map on Demand" project has been providing a

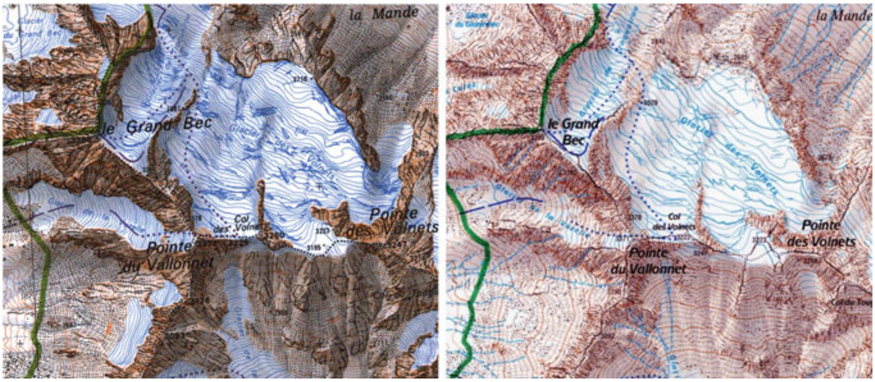

Fig. 11.6 Mountain representation: (*left*) current map obtained manually in the past; (*right*) test with numeric representation

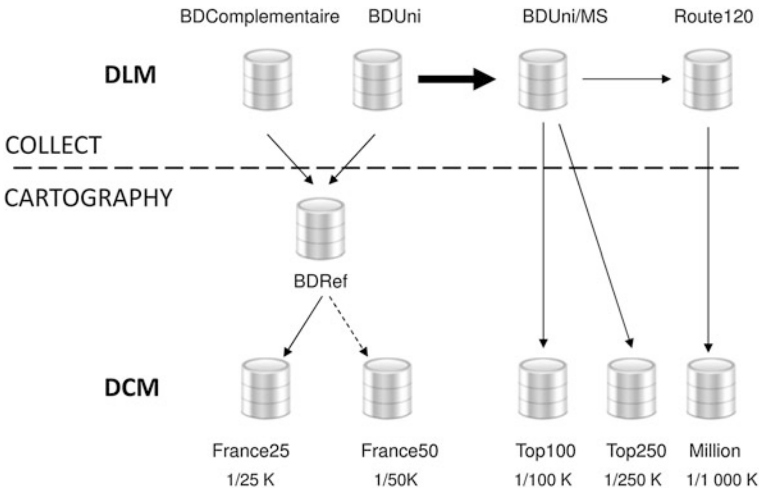

Fig. 11.7 Derivation processes at IGN in 2013 (*thin arrows*) and future work (*thick arrow*)

fully automatic process to map the BDUni for the whole of France within one week for an express product. Such an express mapping product is of a lower quality than Type 2010, particularly in terms of generalisation, but it is more up to date and offers the possibility of selecting different options of colour, and so a better customisation. After new investment into automating further the generalisation processes, we can assume that this product will be gradually improved and will obtain the same quality as Type 2010 currently obtained with semi-automatic processes.

In the future, IGN will continue to work on derivation processes. After considerable effort in cartographic derivation (partly reported in this section and summarised in Fig. 11.7), the next challenge will be to introduce the derivation between two DLMs at different level of detail. In 2010, IGN migrated the BDCarto® into the same DBMS as the BDUni: PostgreSQL. The product

BDCarto® was renamed in BDUni/MS (MS for Medium Scale) to distinguish from BDUni that corresponds to large scale. But, IGN still has to manage two different update cycles, with two separate collect processes, for these two DLMs. To further reduce production costs, it would be necessary to derive BDUni/MS from BDUni and then to be able to propagate the updates. This would be the next step in the evolution of managing derivation problems in production.

11.4 Producing Digital Cartographic Models at Swisstopo

Dominik Käuferle

«That day all the church bells of Switzerland rang in my heart» stated Professor Eduard Imhof on 21 June 1935 after the federal council had approved the law required for the creation of new National Maps of Switzerland. Now, in 2013—the year of swisstopo's 175 years anniversary—a new era begins as swisstopo completely modernises the National Map series. The new National Maps are based on Digital Cartographic Models (DCM), in seamless vector databases. The new map data are fully GIS based and are fit as well for mobile and online services, while offering the aesthetics and readability of high-quality printed maps. Automatic generalisation plays a key role in this new production environment.

11.4.1 A New Data and Database Infrastructure

Thirteen years ago, swisstopo decided to investigate new ways of producing its base data. Two projects were initiated. One to build a central topographic 3D database, the Topographic Landscape Model (TLM), and another to derive cartographic databases from the TLM, the Digital Cartographic Models (DCM). In both projects automation was key to producing a huge amount of data in a short time. The concepts were completed in 2004. The realisation began in 2005. The key components were:

- TLM (Topographic Landscape Model)
  - seamless topographic database for the whole territory of Switzerland
  - photogrammetric restitution of aerial photos
  - vector data model with more than 100 basic topographic object types, combined with a high resolution digital terrain model
  - 3D geometries with accuracies of ±1 m
  - no cartographic generalisation

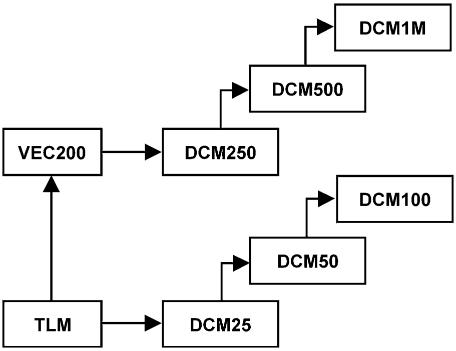

Fig. 11.8 Overview over the new TLM/DCM data infrastructure

- DCM (Digital Cartographic Models)
  - one seamless database per map scale (DCM25, DCM50, …)
  - derivation from TLM
  - a requirement to preserve or exceed the quality as compared to the existing National Map series, which is to be achieved with the same or less human production capacity
  - 2D vector data model combined with raster layers (relief, rock and scree representation)
  - combination of automatic and manual generalisation
  - unique identifiers and feature links between DCMs and TLM to maintain database consistency and facilitate incremental updates
  - cartographic symbols adapted to TLM classifications
  - new fonts

The TLM and the large-scale maps 1:25,000, 1:50,000 and 1:100,000 are updated in a 6 years cycle. The smaller scales 1:200,000, 1:300,000, 1:500,000 and 1:1,000,000 are derived from an intermediate landscape model VECTOR200, which has a yearly update cycle.

Figure 11.8 gives an overview of the new data infrastructure and the conversions (black arrows) from one database to another. In short it can be described as a combined star and ladder approach, with two stars (TLM and VECTOR200) and two consecutive ladders (large-scale DCMs and small-scale DCMs). An attentive observer of the project gave us the feedback that the ladder approach might lead to inconsistencies in smaller scales due to error propagation. To avoid such problems, swisstopo will compare the child data (e.g. DCM50) during production not only with its parent data (e.g. DCM25), but also with the TLM. Based on the feature links, such comparisons can be automated to some degree.

11.4.2 The Role of Automatic Generalisation

In the workflow of the current National Map series, only recognised changes from reality initiate map updates. In contrast, producing DCMs means that all objects, whether they have to be updated or created for the first time, rely on the TLM. Generalisation is necessary in order to receive the required cartographic result. Findings at the concept phase of the project showed that creating the DCM would imply about a fivefold workload compared with the one needed for the manual work of a complete map update. This means, that at least 80 % of the work needs to be done automatically via available cartographic resources. Since a considerable amount of work is cartographic generalisation, automatic generalisation plays a key role. Without automatic generalisation, it will not be possible to produce the DCMs in a reasonable timeframe. The DCM production is done in map sheet perimeters and consists of four consecutive steps:

1. completion of TLM including semantic quality control;
2. model generalisation of TLM to a raw DCM, also referred to as cartographic reference model;
3. automatic generalisation of DCM vectors and automatic scree representation;
4. manual completion of the DCM including cartographic quality control (vector editing, raster editing, labelling, integration into the seamless database).

The automatic generalisation of DCM vectors in Step 3 is performed by «SysDab». This automatic generalisation system was built by *Axes Systems AG* on behalf of our project and is based on their software *axpand ng*. SysDab performs the main portion of the automatic generalisation work. The system has a client server architecture, in which data are stored in an Oracle database. It offers a variety of generalisation functions that perform specific vector generalisation tasks while maintaining topology, such as:

- building generalisation and alignment;
- parallel adjustment of line and polygon segments;
- line, area and point displacement;
- snapping polygon segments to lines and vice versa;
- point typification;
- merge functions.

The functions can be configured to reach specific generalisation goals. The core of the system is a workflow manager that allows all the tasks to be placed together in a customised workflow. The principles of this workflow management were described by Petzold et al. (2006).

As the devil is in the detail, several approaches and iterations were necessary to satisfy swisstopo's high-quality cartographic requirements. Finally, generalisation rates of 78 to 84 % were successfully generated in DCM25 samples. This means, that only 16–22 % of the features in the test map needed manual editing on their geometries. Even though these are good results, there is potential for improvements.

The processing was performed on a server containing 4 CPUs with 24 cores. Processing time ranged between 4 and 5 h per map sheet. The workflows have been redesigned and are constantly fine-tuned in order to reach even better generalisation results.

SysDab and the swisstopo workflow are built for incremental updates and generalisation and allows us to generalise the data several times. The system keeps track of the generalisation for each feature and maintains their history and relationships to its predecessors. A customised relevance check determines if changes in the input or output are relevant enough to be processed or exported.

If one defines *generalisation* as being synonymous with *abstraction*, a wider range of automatic generalisation tasks can be identified in the DCM production chain. The TLM production (Step 1) includes an automatic vegetation extraction. Digital surface models and infrared orthophotos are used to extract trees automatically, which are then generalised (abstracted) to forest polygons or 3D points for single trees in the TLM. Step 2 is performed fully automatically using *FME* by *Safe Software*. One interesting example is the automatic scree representation in Step 3. Scree is represented by small irregular black polygons (scree points) within a scree area. The size and distribution of the scree points depends on the relief shading—the darker the relief is, the larger and denser the scree points are. The software *screepainter,* engineered by Jenny on behalf of the project (Jenny et al. 2010), performs this rather complicated task in a fully automated manner. In Step 4, the *ArcGIS* based cartographic system «Genius-DB» developed by *Esri Switzerland* on behalf of our project, includes a large variety of automated tools that efficiently assist the cartographer in implementing sophisticated cartographic rules.

Even if all these automation tasks could be done without a database infrastructure, it helps a great deal to integrate all the different tasks into a robust and manageable production chain. Key elements are central data dictionaries, which are repositories for the rules, configurations and process steps performed on the different data models.

11.4.3 The Benefits of DCM's

DCMs are part of the national geodata infrastructure (NGDI) of Switzerland according to the Federal Act on Geoinformation (The Federal Authorities of the Swiss Confederation 2007). The DCMs will be included in the federal geoportal http://map.geo.admin.ch/ and in swisstopo services, such as the new «journey through time» service (Swisstopo 2013). DCMs are static datasets containing well-structured vector data. This offers the following advantageous possibilities:

- apply cartographic templates (e.g. colours);
- link cartographic data with third-party data;
- process and provide data independent of the map sheet lines;
- deploy the detailed structure by layer or object types;
- (long-term) monitoring of landscape processes at different scales.

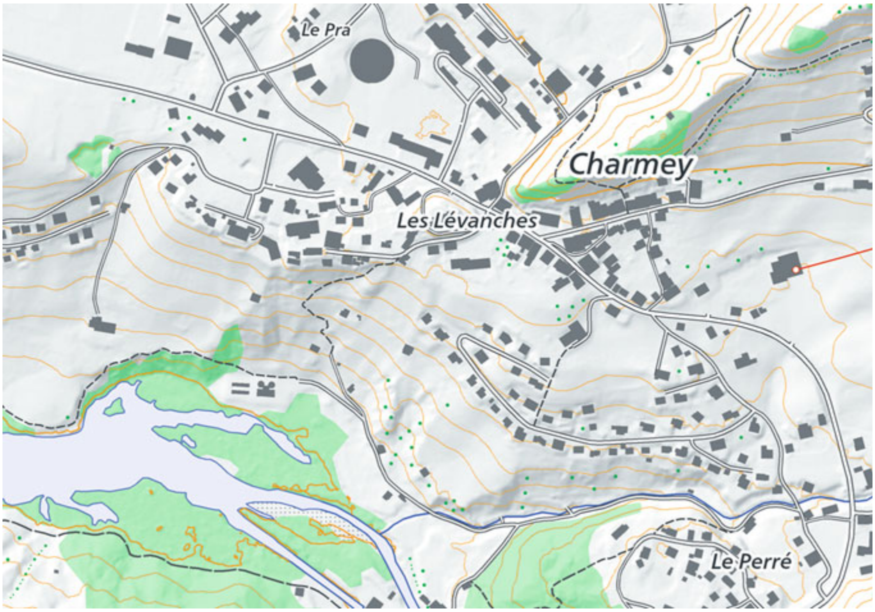

Fig. 11.9 Cut-out of a TLM extract, 1:10,000 (no generalisation applied, automatically symbolised)

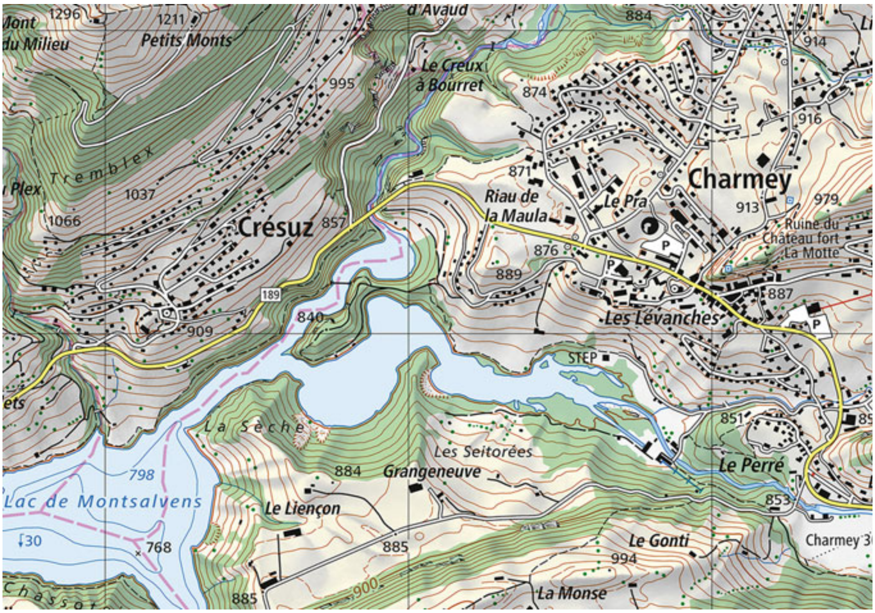

Fig. 11.10 Cut-out of a finished DCM25, 1:25,000 (generalisation applied, test sample)

These opportunities will open several doors for the use of swisstopo map data. There will be new and 'in-house' products, and tailored maps produced by external individuals and companies. Production of the DCM25 (Figs. 11.9 and 11.10) commenced in March 2013, but there remains a lot of work to do. In the sense of Eduard Imhof: our hearts are beating fast at the moment. The completion of all DCM's is planned to be accomplished within the next 8 years.

11.5 Automatic Map Derivation at Ordnance Survey GB

Nicolas Regnauld

11.5.1 Introduction

Ordnance Survey has recently started a large programme of work to change the way products are created. This was triggered by a clear change in what Ordnance Survey's customers want, as well as a need to reduce production costs. The programme is called the Multi-resolution Data Programme (MRDP), and aims at deriving products from Ordnance Survey's large scale topographic database, in a flexible and efficient manner. It also seeks to bring greater consistency across the range of products on offer. Section 11.5.2 describes the strategy developed by Ordnance Survey to derive products in the future. This includes a high level description of the database architecture used. Section 11.5.3 briefly describes the technologies used in the system. Section 11.5.4 presents OS VectorMap® District v1.0, the first product derived automatically by MRDP from Ordnance Survey large scale data. Finally, Sect. 11.5.5 discusses future plans, related to incremental updating.

11.5.2 Product Derivation Strategy

The architecture designed for this system is based on a multi-resolution database. We call it multi-resolution instead of multi-representation, as the term representation may be confused with cartographic representation, while we want this database to stay as independent as possible to the representation used in specific products. However, the concept is the same, and the multi-representation database has been widely studied (Balley et al. 2004). This architecture follows the DLM/DCM (Digital Landscape/Cartographic Models) principles, first presented in (Grünreich 1985) and now widely adopted (Trévisan 2004; Bobzien et al. 2007). The schema in Fig. 11.11 shows a simplified view of the architecture, focusing on

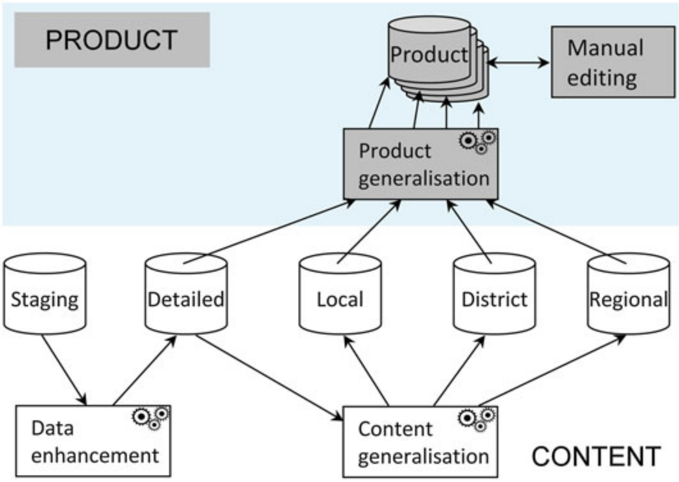

Fig. 11.11 Simplified architecture used by MRDP

how the database is organised. The staging database is a mirror of the maintenance database that contains all the large scale topographic data. It is kept separate to prevent the processes involved with deriving products from interfering with those involved with updating the main database. This database is then used as a source for the data enrichment processes, to make implicit information explicit (such as creating urban extents, deriving networks from topographic features, etc.). The results of enrichments are stored in the detailed content database. This is then generalised to populate the lower resolution databases. Each content database is then used to populate specific product databases. Content databases at each resolution (detailed, local, district, regional) are used as a single source for different products at similar resolution. The intention is to reuse the same content data for efficiency and consistency purposes.

All databases will be updated incrementally. The change data will arrive in the staging content database, and be generalised to update all the lower resolution content databases. The regime of update for the products is product dependent. The change will be processed in chunks—something we call clusters. A cluster is a set of features affected by change that need to be processed together. These are dynamically constructed, sometimes as an aggregation of smaller static clusters (such as partitions formed by road lines). This aggregation ensures that the change will be encapsulated inside a cluster and not across its boundaries. There will be several types of cluster used to propagate different types of change. This change only update mechanism is only theoretical at the time or writing, and has not been implemented beyond a quick proof of concept.

This architecture results in a lot of data redundancy. For example if the only difference between two resolutions for a particular theme is that small features have been eliminated at the lower resolution, then all the other features are present in the same form in both databases. This is not a big issue for data that does not need to be manually edited. However, if several products at the same resolution need manual finishing, we try to minimise the data redundancy to minimise the manual effort required. For example, manual editing on a common product database is done before styling them differently in separate ‘child’ product databases. This is only possible if the different styles have certain similarities, such as the size or width of symbols.

11.5.3 Technology Used

A number of technologies are used by the MRDP programme to build the production systems. Some have been chosen because they are already available and widely used in Ordnance Survey. This is the case with Oracle being chosen as the database provider, and ArcGIS as the editing platform. However, we needed to integrate into the system a way of developing bespoke generalisation algorithms, as no existing platform provided the generalisation tools required to meet all the requirements. Ordnance Survey therefore started a competitive tender process at the end of 2009 to choose a platform for developing its own automatic derivation tools. The contract was awarded to 1Spatial in January 2011, allowing Ordnance Survey developers, in partnership with 1Spatial, to develop processes based on Radius Studio and Radius Clarity. These two platforms have been used to perform the data enrichment, content generalisation and product generalisation tasks (Fig. 11.11). More details on these platforms and how they have been integrated for the needs of MRDP can be found in Regnauld et al. (2012).

11.5.4 Results: OS VectorMap District

MRDP has recently completed the production of OS VectorMap District v1.0. Products from the OS VectorMap family have been designed to be easily customisable to create bespoke contextual maps for websites and applications. In addition to the usual raster format, these products are also available in vector format, which allows them to be quickly loaded into a GIS, where each layer can be turned on or off, and styled as required. OS VectorMap is available as part of OS OpenData and current users include local authorities, emergency services, and the insurance industry. It is also an ideal entry-level product for market demographic displays and can be used by anyone for sharing statistics or for neocartography.

Prior to this release, OS VectorMap District had already been made available to the public in the form of alpha and beta versions. While these early versions were produced using research prototypes, the latest one was the first one produced by an MRDP system, in which all the processes from the data extraction, to the publication process have been automated. More information on this system can be

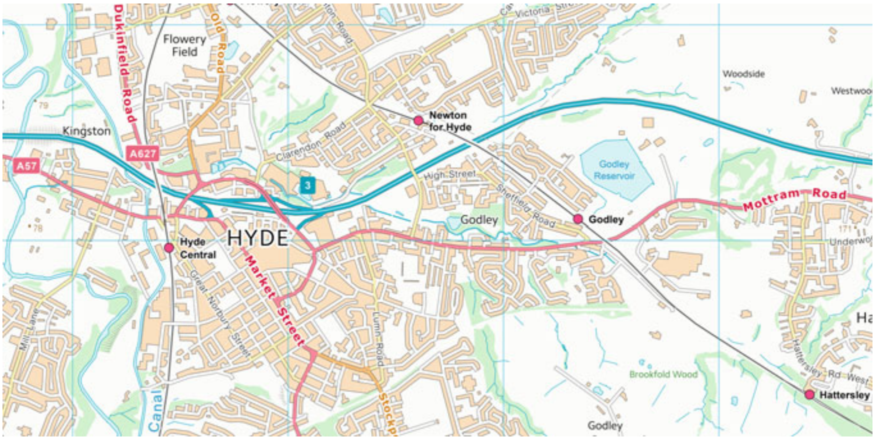

Fig. 11.12 Extract from Os VectorMap District v1.0, west of Manchester. Ordnance Survey © Crown Copyright. All rights reserved

found in Regnauld et al. (2013). Most of the generalisation processes used come from the initial research prototypes, but have been enhanced to ensure they can be maintained and reused by the programme. The detail of the generalisation processes that have been developed for deriving OS VectorMap District can be found in Revell et al. (2011). They include:

- Building generalisation (aggregation, simplification);
- Coastline: creation of a continuous coastline, followed by simplification (Zhou and Jones 2004);
- Vegetation generalisation (aggregation, simplification);
- Railways conflation and generalisation;
- Name extents conflation and filtering.

Version 1.0 of the OS VectorMap District also contains a better depiction of roads, thanks to the use of a process to automatically collapse dual carriageways (Thom 2005). Figure 11.12 shows an extract of OS VectorMap district v1.0.

Figure 11.13 shows an example of use of OS VectorMap District, where only selected layers have been used and combined with height data to produce a map illustrating flood risks in Carlisle.

11.5.5 Future Plans

Now that MRDP has delivered the first system to automatically derive a product from the base data, the next steps are to extend the capabilities of the system to tackle other products which are still maintained using old systems relying heavily on manual editing. For more complex cartographic products, manual editing will

Fig. 11.13 OS VectorMap District combined with a height product to show flood risk in Carlisle. Ordnance Survey © Crown Copyright. All rights reserved

still be required, and one of the challenges will be to add the manual editing process into the workflow. The system will also be extended to perform updates incrementally. While the system has been designed with incremental updating in mind, it has not been implemented in the current system used to produce OS VectorMap District. This was left out in order to stage the development of the system, to avoid tackling too much complexity at once. OS VectorMap District is derived automatically, so incremental updating is not a critical requirement. It will become critical when the system is developed further to include manual updating, as we will want to minimise the loss of previous manual modifications.

11.6 Generalisation Methods Used for the USGS National Map and National Atlas

Lawrence V. Stanislawski, Cynthia A. Brewer and Barbara P. Buttenfield

Original 7.5 minute, 1:24,000 scale (24 k) topographic maps for the United States were handcrafted and field verified paper maps produced and subsequently revised by the United States Geological Survey (USGS) between 1945 and 1992 (USGS 1955; Moore 2011). During this time the 7.5 minute maps formed the foundation for the topographic map series from which smaller-scale generalised versions of the maps were derived. The 24 k maps were manually generalised to smaller scales, such as 1:100,000 (100 k), through photo-reduction of individual maps, panelling of the reduced maps, and re-scribing of features from the reduced colour plates using detailed specifications to adequately reduce map content (USGS 1985, 1993).

With technological advances in GIS, the USGS modernised its mapping program in the early 1990s into a dynamic web-accessible digital version. First, scanned digital versions of the paper maps were made available as digital raster graphic (DRG) files, while vector and raster databases of the primary themes were being compiled or collected. Since then, the USGS has coordinated development and maintenance of *The National Map*, which manages and delivers eight primary geospatial data themes and provides web display and access to these data through *The National Map* Viewer, along with downloadable up-to-date US Topo (24 k only) and historical topographic maps (Sugarbaker and Carswell 2011). Traditional USGS paper maps have been phased out of production.

The USGS Center of Excellence for Geospatial Information Science (CEGIS), in collaboration with the University of Colorado at Boulder and Pennsylvania State University, has coordinated research and development for automated generalisation to support multi-scale display and delivery of *The National Map* and other USGS data. A product of this research is a set of tools that automate methods to generalise the high-resolution (HR) layer of the National Hydrography Dataset (NHD) to 24 k and smaller scales. The NHD is one of the eight themes of *The National Map* and furnishes the surface water component for US Topo maps. Section 11.6.1 describes use of the tools to automatically generalise local resolution NHD to 24 k for display on the 24 k USGS Topo maps. Section 11.6.2 presents automated generalisation of high-resolution National Map data to smaller scales.

11.6.1 Generalisation of Local Resolution (Large Scale) Hydrography for 24 k US Topo Maps

US Topo maps are freely downloadable digital maps that are modelled from the traditional 24 k paper topographic map. The maps are published in Portable Document Format (PDF) with a geospatial extension that is called Georeferenced PDF (GeoPDF®) as patented by TerraGo Technologies. They are mass produced through automated procedures that access and display national GIS databases and through semi-automated annotation and editing methods (Moore 2011; Sugarbaker and Carswell 2011). Currently (2013), the USGS produces about 18,000 24 k US Topo maps per year, with a planned refresh cycle every 3 years (Moore 2013). Although US Topo maps presently have less feature content and aesthetic quality than the traditional paper map series, the maps are superior for several reasons:

they include a high-resolution orthophoto image and the best available, up-to-date data (three year refresh cycle); they are produced in a single modern coordinate system; and they are rapidly produced and freely downloadable through the web which reduces time and expense for users (Moore 2011).

Because of densification efforts implemented by state and local partner organisations, the HR NHD is a multi-scale layer that must be generalised in places from local resolution (1:12,000 or larger) to 24 k for consistent representation on US Topo maps. CEGIS NHD generalisation tools call existing ArcGIS® functions (version 9.3 and later) within customised geoprocessing scripts to generalise subsets of NHD data, which are stored in Esri file geodatabase format. As described in Chap. 6 and elsewhere (Stanislawski and Buttenfield 2011; Buttenfield et al. 2011), the tools automate enrichment, pruning, feature simplification and other generalisation operations, along with validation for subsets of NHD data, while maintaining spatial topology and the NHD model schema for use with other applications.

The NHD generalisation tools execute stratified network thinning operations by separately thinning features in assigned density partitions to target densities (Stanislawski and Savino 2011). More details about stratified network thinning for roads are presented in Chap. 6. Subsequent to network thinning, less prominent polygon features are eliminated by feature type and standards-based minimum area constraints (US EPA and US DOI 1999). Channel hierarchy is assigned to retained flowline features based on network topology and estimated upstream drainage area (Anderson-Tarver et al. 2012). Lastly, retained features are differentially generalised to reduce granularity (e.g. vertex spacing) but preserve geometric characteristics, local density, and texture. The sequence and type of generalisation operations applied in this phase of processing are tailored to local landscape and feature conditions (Stanislawski and Buttenfield 2011; Buttenfield et al. 2011).

Proper network topology and spatial relations among overlapping feature types are maintained according to NHD schema rules. However, supplementary processes are still needed to ensure coincidence among simplified features, such as among simplified coastlines and associated lake or ocean polygon boundaries. Typical subbasins, covering about 25 24 k US Topo maps, require about an hour or less for generalisation processing. Pruning parameters are automatically estimated from archived 24 k benchmark data. Selection of thinning parameters is guided by the Radical Law (Töpfer and Pillewizer 1966) in the absence of benchmarks. Default simplification parameters are available for typical density classes, but require some manual review and adjustment.

Figure 11.14 shows a section of the 24 k US Topo map for Snake Mountain, Vermont with local resolution hydrography and with local resolution hydrography generalised to 24 k using the NHD generalisation tools. Local resolution for Vermont is compiled at 1:5,000 (5 k). Four partitions having local resolution densities of 0.86, 1.69, 2.63, and 3.78 km per square kilometer (km/km^2) were defined for the flowline features within the NHD subbasin associated with the Snake Mountain map. Flowlines within these partitions were separately pruned to densities of 0.54, 1.07, 1.66, and 2.38 km/km^2, respectively, based on the 24 k

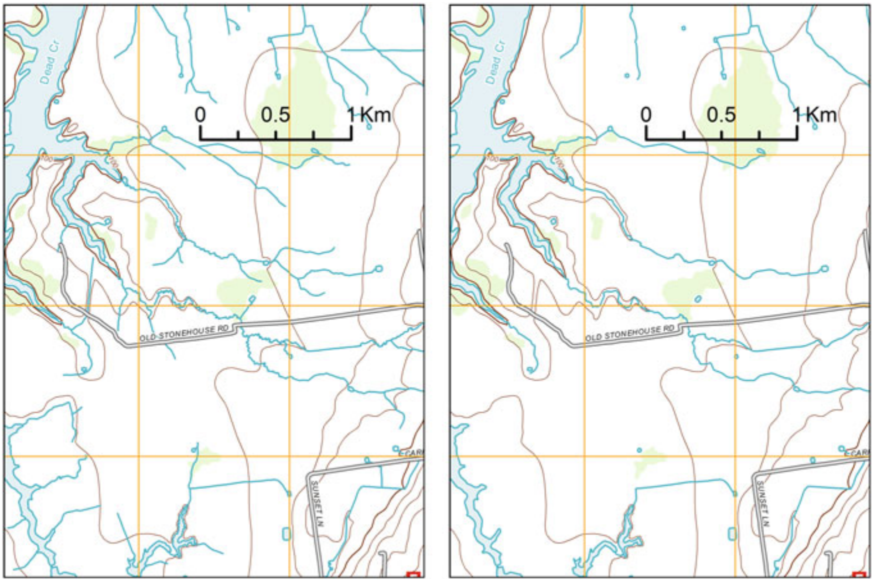

Fig. 11.14 Portion of 1:24,000-scale (24 k) US Topo map of Snake Mountain, Vermont without orthoimage displayed at 1:50,000 with local resolution hydrography (*left*) generalised to 24 k (*right*)

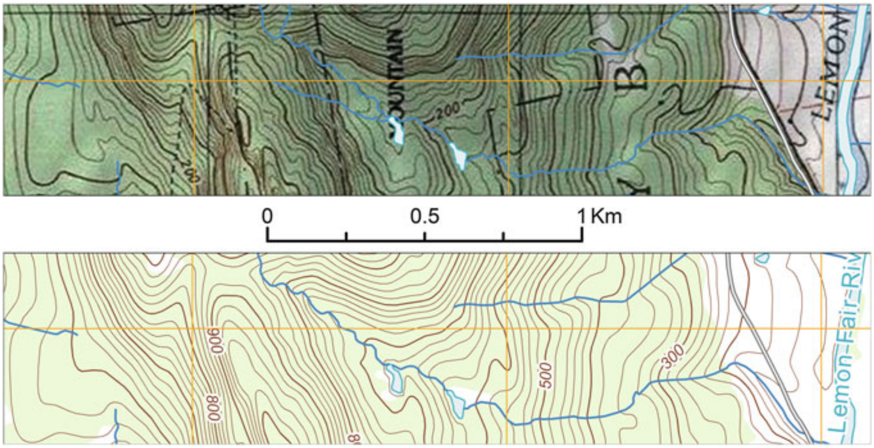

Fig. 11.15 Portion of 1:24,000-scale (24 k) US Topo map of Snake Mountain, Vermont displayed at 24 k without orthoimage. Top panel includes UTM grid and local resolution (1:5,000-scale) hydrography *over* the 1980 digital raster graphic, and the *bottom* panel shows local resolution hydrography generalised to 24 k, along with the map features typical to this product. Stream symbols have been widened to aid visualisation

NHD benchmark. Retained linear and polygonal features were differentially generalised to reduce vertices using the Bend Simplify algorithm (Wang and Muller 1998) available in ArcGIS®.

A closer view of another portion of the Snake Mountain map reveals the effect of line simplification and the quality of integration between the generalised hydrography and contour lines (Fig. 11.15). For each 24 k map, contours are vertically integrated with generalised hydrography through a fully automated hydro-enforced contour generation process (Arundel et al. 2010). This process ensures that stream lines follow valleys and that contours properly pass through double-line streams, without intersecting other waterbodies.

11.6.2 Generalisation for National Atlas (Smaller Scale) Hydrography and Multi-Scale Display

This section briefly reviews generalisation methods applied by the USGS to compile the National Atlas hydrography and presents some generalisation methods and enhanced topographic map designs being tested for multi-scale display of *The National Map* and the National Atlas data.

The 1 M hydrography of the National Atlas was recently produced for the United States by generalising the 100 k NHD through automated and semi-automated feature selection, simplification, and refinement operations. Requirements for this multi-year task were defined for the stream network, waterbodies, coastlines, and associated stream gauges. Feature selection and generalisation criteria, as well as criteria for integration between hydrography features and between hydrography and other associated National Atlas features, such as roads and boundaries, were included. Criteria for selecting which 100 k NHD stream and waterbody features to retain in the 1 M National Atlas were largely based on whether the features exist in one of several benchmark datasets. Integration criteria imposed proper snapping of connected features (e.g. stream ends to coastlines) and displacement between features (e.g. between streams and roads), which was implemented to ensure cartographic clarity. Retained streams and waterbodies were simplified using a 500 m tolerance with the Bend Simplify (Wang and Muller 1998) algorithm in ArcGIS®. Additional procedures adjusted stream densities and feature geometries. For further details see Gary et al. (2010).

The USGS has been working with Pennsylvania State University to redesign the US Topo maps for enhanced clarity over images and for multi-scale display of geospatial data. To do this, generalisation is applied to hydrography, roads, land cover, Geographic Names Information System (GNIS) point features, and elevation themes of *The National Map*. Automated and manual procedures are largely implemented through existing or customised ArcGIS® geoprocessing tools and editing environment. The following describes some of these efforts.

As demonstrated here and by others (Brewer and Buttenfield 2010; Buttenfield et al. 2011), the NHD Generalisation tools can be used to generalise the HR NHD to smaller scale Level-of-Detail (LoD) datasets that can support graphics over a limited range of scales (Cecconi et al. 2002). Figure 11.16 shows three *hydrography* LoDs.

At 50 k (Fig. 11.16a), the 24 k NHD data are shown with minor modifications. Double-line stream areas are replaced by selected primary paths available in the enriched network data, collapsing the area to a line, and small ponds are removed using a minimum area threshold. The 100 k LoD (Fig. 11.16b) has more extensive line pruning, waterbodies removed according to an increased threshold, and differentially simplified lines and polygon boundaries. At 1 M (Fig. 11.16c), National Atlas hydrography displays simplified major rivers and large waterbodies. At all scales, lines are tapered using a small number of width classes based on upstream drainage area for NHD and stream order for National Atlas data, symbolising these attributes in a simple hierarchy.

Local roads are systematically thinned using the ArcGIS® Thin Road Network tool (Sect. 6.5). The tool is applied in a ladder fashion, thinning on minimum lengths incrementing from 500 m to 30 km. Resulting visibility settings are combined into a 34-level attribute. The visibility field controls removal of the least important roads and labels, guides priorities by which the Esri Maplex Label Engine places text labels, and enables emphasis of labels for roads that are major thoroughfares within a mesh of urban roads (Brewer et al. 2013). Overlaid on this hierarchy of text labels and thinned local roads is a network of interstate and state routes, with their hierarchy established by separate road-class attributes. Comparison of Fig. 11.16a–c shows fewer local roads (light with a brown casing) that make connections across the highway network as scale decreases.

A basic landscape characterisation was produced using the National *Land Cover* Dataset (NLCD). These data were processed by raster operators to achieve a generalised, smoothed, and interpolated representation at 24 k, resampling this layer repeatedly for representations at smaller scales. Interpolation was necessary because 30 m-resolution is too coarse for 24 k display. For classified land cover data, random pixels at raster region boundaries were combined with proximal pixel values to create an "airbrushed" effect for 2 m-pixel upsampled land cover. For canopy and impervious surface data, ladder-style bilinear upsampling from 30 to 2 m pixels interpolated graded representations. Hues with similar lightness across all land cover categories produce visually vague boundaries, and these colours grade to white for high percent impervious surface and to dark greens for high percent canopy cover (Fig. 11.16a, b).

Populated places from GNIS are categorised into text label size categories based on related decennial census data from the U.S. Census Bureau. Low-population places are removed as display scale decreases. Topographic maps provide an important resource for emergency response in the United States. A selection of *structure types* from GNIS point features is shown in Fig. 11.16. Points for fire station, police, hospital, airport, and school locations are selected using feature codes, with additional queries on name content to remove point subsets. Text labels and symbols for less important structure types are removed at smaller scales; for example, schools are removed from Fig. 11.16b, c.

Terrain form is also generalised for smaller scales (Fig. 11.16). At largest scales, the 10 m-pixel (1/3 arc-second) National Elevation Dataset (NED) is smoothed once using a low-pass filter with a 3×3-pixel kernel, and then similarly smoothed

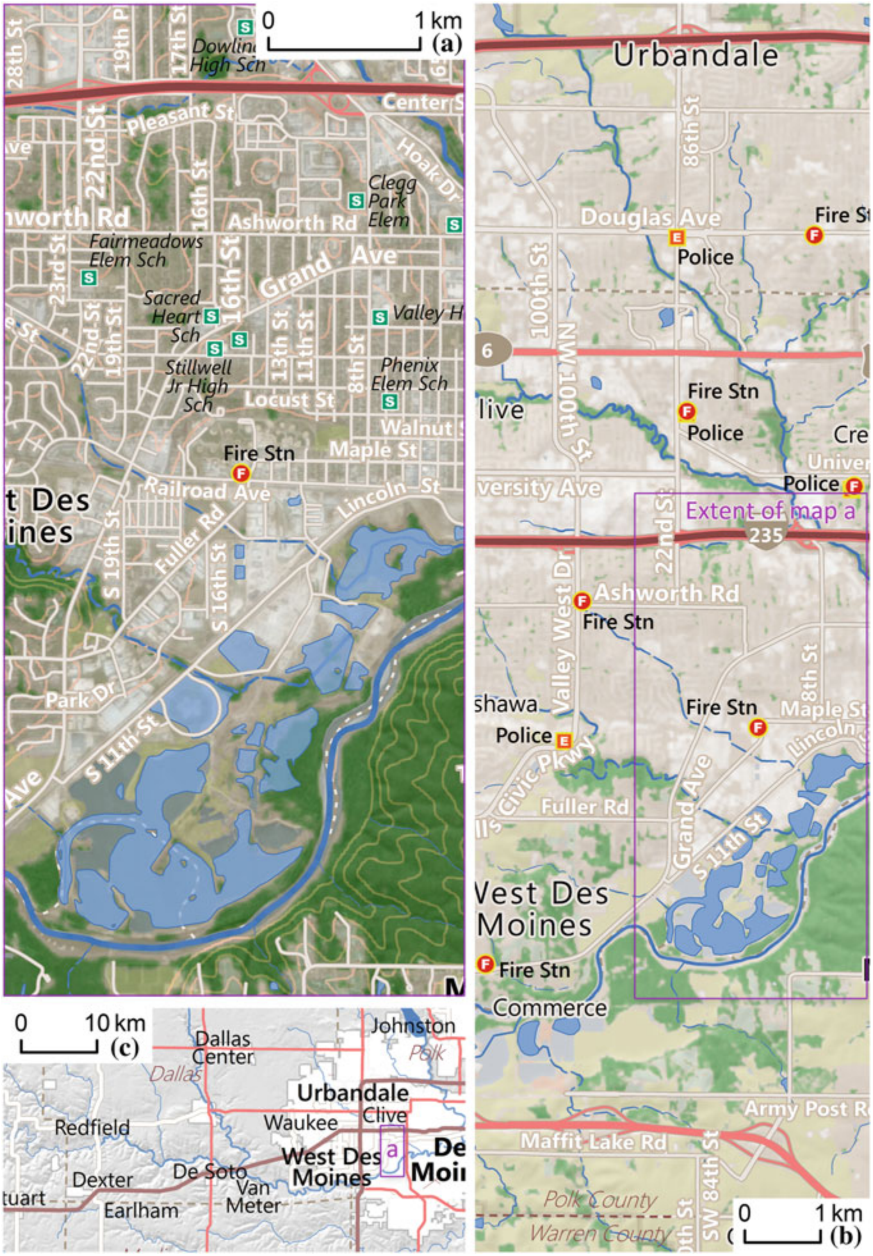

Fig. 11.16 Three example scales that anchor the continuous-scale topographic map design in development, showing a portion of West Des Moines, Iowa. Figures show generalised roads and hydrography: **a** at 1:50,000 with land cover, shaded relief, contours, and an orthoimage background; **b** at 1:100,000 with landcover and shaded relief; **c** at 1:1,000,000 with *shaded* relief only. These landscape layers can each be set off or on, and the design is intended to function for all 16 possible combinations of the four

ten times with a 5 × 5 kernel for middle scales. The 30 m (1 arc-second) NED is smoothed for a base layer at small scales. Smoothed NED data are used to produce five-direction hillshades, a curvature layer (to selectively highlight and generalise ridge and valley lines), and contours for use at large, middle, and smaller scales.

11.6.3 Future Plans

Much work remains for complete modernisation of the USGS mapping program. Research has developed strategies and tools to automate generalisation of some of *The National Map* data themes to LoD datasets required for multi-scale display. These procedures must be refined with additional automation and integration techniques to ensure adequate feature snapping between and within data themes. Further methods may enhance 24 k US Topo maps with additional and revised feature content. Plans are being designed to implement automated generalisation to expand US Topo maps with medium and small scale products.

11.7 Generalisation in Production at Kadaster NL

Jantien Stoter, Vincent van Altena, Ron Nijhuis and Marc Post

11.7.1 Introduction

In 2010 the Netherlands' Kadaster, the national mapping agency, started a feasibility study to examine how automated generalisation techniques could be introduced into its map production environment. The study led to a fully automated workflow to produce 1:50 k maps, which will come on line from 2013. The following sections describe the Kadaster generalisation approach. Section 11.7.2 describes the main characteristics of the approach, Sect. 11.7.3 describes implementation details and Sect. 11.7.4 closes with results, findings and future plans.

The research that resulted in the automated generalisation workflow is described in Stoter et al. (2013).

11.7.2 Main Characteristics of the Kadaster Approach

The feasibility study on automated generalisation first focused on the workflow from 1:10 k data to a 1:50 k map, to be extended to other scales once successful results had been achieved. From the beginning it was clear, that the automated generalisation workflow should not aim to replicate the existing map, even though the new automatically generated 1:50 k map will replace the existing map. There were several reasons for this. Firstly, starting from legacy topographic products—more than 60 years ago—might overemphasis past requirements and may ignore new opportunities of multiscale topographic information. For example, topographic information is used in more applications by a wider public than ever before and the user may prefer up-to-date maps over maps that meet all traditional cartographic principles (though the results should still be of an acceptable quality). Furthermore, automating a previously interactive process, which was designed in a previous technical and organisational context, can be very complicated (Foerster et al. 2010; Stoter et al. 2009b). Another reason to reconsider existing map requirements is that the time-saving aspect of automation makes the process well suited to the production of multiple products that meet different demands. The consequence of this is that the requirements of automatically generating multi-products for a certain scale differ from the requirements of the existing single product. A few other aspects further refined the scope of the Kadaster approach:

- The focus is on producing a map. We acknowledge that users may require mid- and small-scale data for specific themes to be applied in spatial computations. Examples are water and road networks. For those themes separate data products will be considered;
- The topographic data of Kadaster forms a partition of space at every scale. This brings specific challenges, since every object that is removed needs to be replaced by something else. In addition, the displacement process always causes other objects to be displaced as well;
- Generalisation without any interaction is the best guarantee for efficiency and consistency and the only way to produce multiple on-demand products. Therefore interactive improvements of the generalised results is not allowed;
- The most straightforward way for updates in automated generalisation is to completely replace the old version map (Regnauld 2011). At the current time, we do not maintain links between the objects at the different scale levels. If these links are required, this will be part of a subsequent study.

One of the main challenges was how to redefine specifications for automated generalisation by modifying existing guidelines while ensuring that users requirements are met. Firstly, we generated an initial 1:50 k map in a semi-automatic manner by extending the work of Stoter et al. (2009a, 2012). The aim of this first step was to see how much automation we can achieve with currently available tools and some in house developed algorithms. This work implemented existing generalisation guidelines for interactive generalisation in an automated process and improved the implementation by evaluating intermediate results.

The initial map was sent to a selection of key customers of the existing 1:50 k map to test the main principles and assumptions. Based on these insights the process was improved and refined and implemented as one integrated workflow. In the next stage the evaluation and improvement process was repeated, by asking more costumers as well as a panel of typical individual users to assess the resulting map in more detail and for different areas. Based on those evaluations and iterative testing, the optimal sequence of steps was determined as well as the most appropriate algorithms and parameter values for each step. Internal validation via cartographic experts was also undertaken. After two years, the feasibility study resulted in a fully automated generalisation process implemented as an integrated workflow in 2013.

11.7.3 Implementation Details

The following four subsections describe the software, the pre-processing of the data, the implemented automated generalisation workflow and finally the approach that was applied to automatically generalise the whole country.

11.7.3.1 Software and Technology

For the implementation we used a mixture of standard ArcGIS tools, self-developed tools within Python and a series of FME tools. ArcGIS contains some specialised generalisation tools for collapsing two lanes of a road into a single road line, displacing symbolised geometries, simplifying symbolised buildings and thinning networks (Punt and Watkins 2010).

The complete generalisation workflow is implemented within the Model builder tool of ArcGIS. The workflow consists of three main models, consisting of about 200 sub models that are responsible for each specific generalisation problem that we need to solve in the process.

11.7.3.2 Pre-Processing the Data

Since the goal is 100 % automation, the process should cover as many generalisation aspects as possible. This is accomplished by either improving the process step-by-step, or—if that did not work—by improving and enriching the source data. In addition to correcting errors, enrichment of the source data is undertaken in two ways. Either external data sources are used or the required knowledge is made explicit via cartometric analysis. Examples of enrichment of the input data are determining urban extents by defining areas with high density of buildings (i.e. higher than 10 %) and attributing TOP10NL road segments with information on highway exits so that these can be treated in a different manner in the generalisation process from the road network (derived from the TOP10NL roads).

11.7.3.3 Implemented Workflow

The implemented generalisation workflow consists of the following steps:

1. Model generalisation to reduce the data that has to be visualised;
2. Symbolisation of the data;
3. Graphic generalisation to solve cartographic conflicts.

11.7.3.4 Model Generalisation

The model generalisation process is the largest part of the process. Model generalisation is not only conversion of geometric objects to the lower density and structure of the TOP50 model, but also translation and reclassification of attributes to the TOP50 model. Examples of the operations performed in this part of the process are:

- Convert TOP10NL road centrelines (representing single lanes) into TOP50 centrelines (a complete road). The TOP50 centrelines are generated with an ArcGIS algorithm that merges two lanes of a single road into one centreline.
- Convert areas with many buildings (i.e. coverage >10 % which has been experimentally determined as an optimal threshold) to built-up areas and remove the original building objects in those areas. This conversion is only done for areas located within urban area (computed) and if they are not part of an industrial estate (external data). Important buildings such as schools, hospitals, and churches are kept as separate point symbols. This functional information is added in TOP10NL.
- Prune the road network where there is insufficient space to visualise all roads. This is not straightforward as can be seen in Chaudhry and Mackaness (2005), Thom (2007) and Thompson and Richardson (1999). Specifically because TOP10NL data does not contain many attributes for pruning. The pruning of the road network consists of several steps:

    a. Cycle paths parallel and touching roads are deleted; the others (i.e. free cycle paths) are maintained.
    b. Access roads to buildings in rural areas are detected and selected. The automated detection finds access roads "if buildings are located within 200 m of the end of the road AND no other roads are located in the neighbourhood (within 200 m) of the building".
    c. The remaining road network is pruned by the 'thin road network algorithm' available in ArcGIS. This algorithm retains connectivity and the general character while using a hierarchy of the relative importance and the minimum length of a road.

- Prune hydrographic network. First by removing small water bodies parallel and touching roads and then by applying the thin road network algorithm on the linear waterways. Lake shorelines and other water area boundaries are added in this process to maintain connectivity. Thinning is done with the road network algorithm because the Dutch water network is almost completely anthropogenic and its structure resembles a road network rather than a natural water network. Therefore algorithms available for pruning natural hydrographic networks, as studied for example in Stanislawski and Savino (2011) are less appropriate.

The last step in the model generalisation process is the reduction of the data by filtering the vertices of all linear geometries (linear objects and polygon boundaries) by applying the Douglas-Peucker algorithm (parameter value 1 m).

After the model generalisation, the symbolisation process assigns symbols to all geometries, as they should appear on the map. The symbolisation sometimes results in objects that appear larger on the map than they are in reality. The symbolisation therefore governs the next graphic generalisation process that solves the cartographic conflicts. In this process basic symbols are used which precisely correspond to the shape and outline of features, but which lack cartographic refinement. Sophisticated symbolisation and cartographic enhancement to enhance legibility (such as aligning symbols with features) are postponed to a later stage in the process.

11.7.3.5 Graphic Generalisation

The graphic generalisation process consists of the following steps:

1. Generalise (i.e. select, simplify and displace) the buildings that remain after the data generalisation process to avoid overlap and to meet a minimum building size.
2. Displace the linear objects (roads, water) and boundaries of symbolised water and terrain objects as well as all other point and linear objects (i.e. administrative boundaries, height contours, engineering constructs) with an algorithm that displaces symbolised objects and reshapes them in order to avoid overlap. The displacement algorithm uses a hierarchy of object types.
3. Rebuild terrain and water polygon-objects from the displaced boundaries and assign the former codes to the new areas by using left/right information of the boundaries.

11.7.3.6 Country Wide Coverage

To be able to generalise a map for the whole of The Netherlands, we apply the workflow on about 400 generated partitions, obtained from linear objects that must never be displaced, which are the main roads such as highways. Besides some global operations that are applied for the whole country (such as creating and simplifying the power line network), the workflow is applied per partition and partitions are connected afterwards. Because vertices of objects at and near partition boundaries are prohibited from moving in the displacement process, the objects at neighbouring partitions still fit together after generalisation. To process the generalisation of the 1:50 k map from 1:10 k source data for the whole country in a reasonable time we make use of the multiprocessing capabilities within Python. This allows us to process six partitions in parallel on each of the six available systems.

11.7.4 Results, Findings and Future Plans

Figure 11.17 shows the 1:50 k map that is generalised from the TOP10NL data using the fully automated workflow described above. Although our users could not identify significant differences with the traditional version, differences do exist because of the changed map specifications. For example the interactive generalisation guidelines prescribe that detached houses may never be converted into built-up areas. It was impossible to always meet this rule, because detached houses are often enlarged to meet the minimum size requirement (15 × 15 m). Most building blocks containing detached houses are covered with detached houses on both sides, so the widths of the building blocks should be at least 40 m in order to accommodate enlarged detached houses at a minimum distance of 10 m (required for readability). At the same time streets in TOP50 were symbolised with a line width of 20 m, which is wider than streets widths in reality. To solve this, we decided to always convert detached houses in built-up areas, if the building density threshold was exceeded.

Another example of a specification that was modified related to ditches which were not treated as separate objects but as ordinary terrain boundaries, because of their limited importance and because their visualisation is identical to terrain boundaries. All water with linear geometries in TOP10NL (i.e. smaller than 2 m) parallel to roads is eliminated in the new map to save space as we regard them as unimportant. We allow dikes to be displaced. This is because dikes are not available as single objects in TOP10NL, but only visible by hatches on the map. User consultation showed the unimportance of this rule in combination. Given the difficulty of identifying dikes, it was decided to ignore this guideline. Dikes are still present on the 1:50 k map (with hatches), but they may have been moved to make space for other objects.

Based on the results and user evaluations the Kadaster decided that a fully automated generalisation workflow was the most sustainable workflow for the future as well as the only way to produce products on-demand. The model will be implemented from early 2013. From that moment the new maps will replace the existing 1:50 k map product. With thirty-six parallel processors run via Python we are able to perform the core generalisation process for the whole of the Netherlands in approximately 50 h. The aim is to achieve a three week turnaround including pre-processing, generalisation and visualisation.

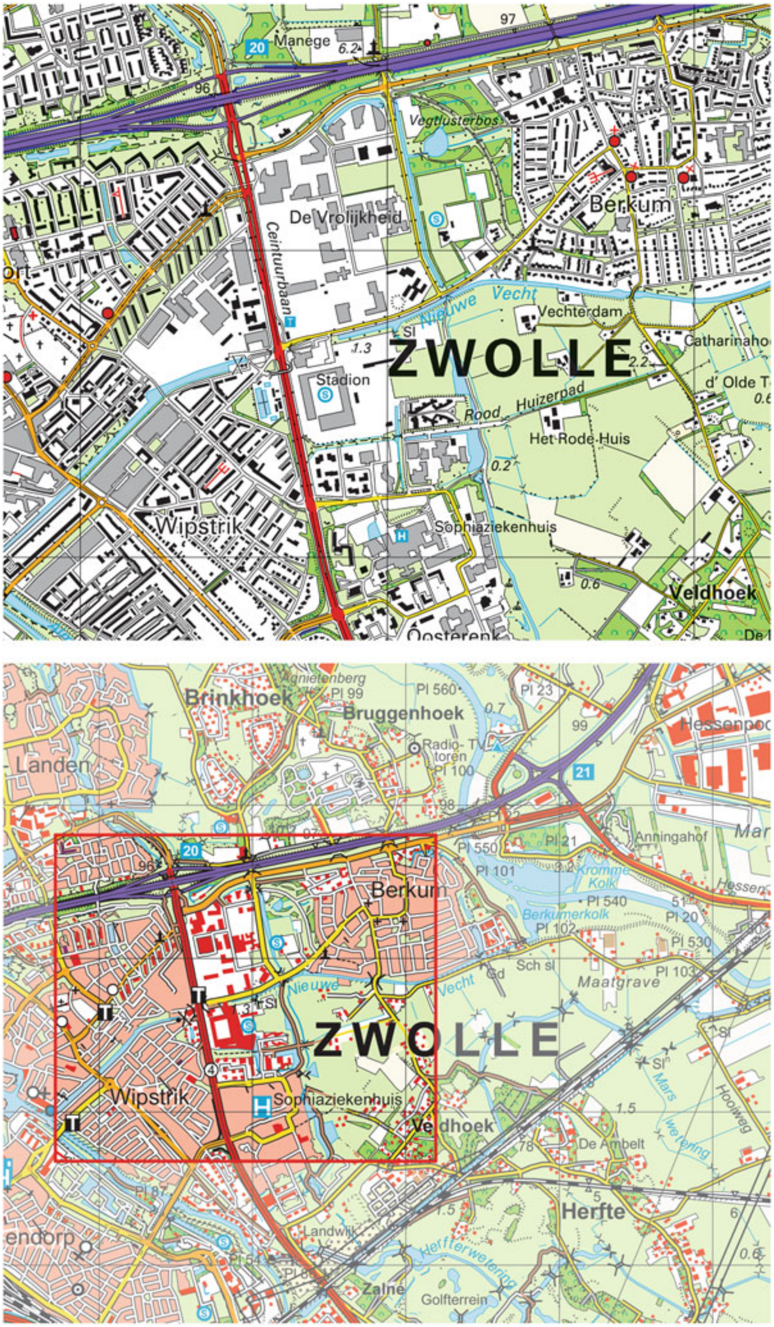

Fig. 11.17 *Top* source data (1:10 k data). *Bottom* generalised map (1:50 k), obtained fully automatically. Both are displayed at smaller scales

Based on the experiences with the new 1:50 k product, the automated generalisation approach will be extended to the 1:100 k map and to on-demand products, such as the backdrop map at multiple scales for the national geo-portal. These products will all be derived directly from TOP10NL, and therefore TOP10NL will become even more important as the base data. Future work will also study the derivation of TOP10NL from the countrywide large-scale data set (approximately 1:500), which is collected by organisations other than Kadaster (mainly municipalities) according to an information model that was established in 2012, called IMGeo (Information Model Geography). Beside scale differences this derivation process needs to cover different perspectives on topographic data, which will be an interesting challenge.

11.8 AdV-Project "ATKIS: Generalisation"—Map Production of DTK50 and DTK100 at LGL in Baden-Württemberg

Sabine Urbanke and Antje Wiedemann

The "Working Committee of the Surveying Authorities of the States of the Federal Republic of Germany" (AdV) established the "Official Topographic-Cartographic Information System" (ATKIS) which was established to map the landscape of all "German Länder" in 1998. Subsequently the automatic derivation of Digital Topographic Maps (DTK) from this database the Digital Base Landscape Model (Basis-DLM) has become increasingly important. For that reason AdV has developed the "ATKIS-Generalisation" project in 2002. The overall objective of the project was to develop software tools to enable the production of Topographic Maps (TK) within a highly automated process, based on a two-stage approach: Model Generalisation (MG) and Automatic Cartographic Generalisation (AKG). Currently twelve "German Länder" are collaborating together with commercial partners to address this task.

11.8.1 Map Production Workflow

In the Basis-DLM the real world objects are modeled as point, line and area vector objects with additional information, such as names and regions. The landscape is structured according to the ATKIS Object Catalogue (OK) in objects and differentiated by attributes. The Basis-DLM is a database with following features:

- 2D georeference,
- high positional accuracy (±3 m), digitised from Orthophotos with a reference scale of 10 k,

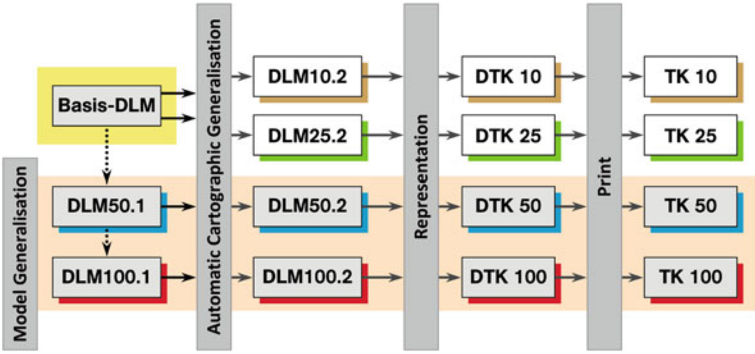

Fig. 11.18 Graphical representation of workflow

- covering the complete national area.

The Basis-DLM is the only database which is continually updated for the entire country. All other DLM are derived by automated processes from this Basis-DLM (Fig. 11.18). The map data are generated within a multi-stage process. The process starts with Generalisation of Basis-DLM data using Model Generalisation producing DLM50.1 (DLM100.1), followed by Automatic Cartographic Generalisation producing DLM50.2 (DLM100.2). The result is subsequently represented, according to the requirements of the style sheet catalogues SK50 (SK100) for DTK50 (DTK100). All the data are stored in a database and can be provided to customers in a specially defined format (NAS-format) and are printed as TK50 (TK100). The generalisation processes are entirely rule-based and can be run independently and separately from each other.

11.8.2 Generalisation Tool: Model Generalisation

The Model Generalisation (MG) is a fully automated process for converting a detailed model (e.g. Basis-DLM) into another, lower density, positionally accurate model (e.g. DLM50.1). The tool has been developed by the company 1Spatial (UK) based on the product Lamps2. The individual generalisation processes are independent of graphic design requirements. They are divided into the steps of semantic and geometric generalisation. The biggest advantage of the Model Generalisation process is to reduce the data volume, which is currently around 33 %.

Semantic Generalisation Step: The transition from the Basis-DLM to the DLM50.1 is defined in "transfer rules". This means the object types and their attributes which are defined in the object catalogue of DLM50, are transferred directly or are changed into a different object type (e.g. through elimination or aggregation). Selection criteria define when an object is to be transferred and what its minimum size must be in order to transfer it to the new database. The tool

Fig. 11.19 Results of model generalisation

checks these conditions and eliminates or merges objects or attributes automatically so they are semantically correct. A similarity matrix regulates and evaluates the semantic combination of the object types.

Geometric Generalisation Step: In a geometry type change (e.g. area to line or area to point) the topological relationships between neighbouring objects must be restored, so that no topological gaps arise in the data. This is done by resolving these objects and creating new geometries according to the modelling rules. Using Douglas-Peucker algorithm, the vertices of linear objects are generalised geometrically, resulting in a thinning of the number of points. This contributes significantly to a reduction in the volume of data. Relations between origin and target objects are kept during the transfer from Basis-DLM to DLM50.1 (DLM100.1).

The production of the DLM50.1 dataset for the whole country takes 95 h. To produce the DLM100.1 dataset the Model Generalisation process is started again with a modified parameter set, running for 29 h (Fig. 11.19).

Number of objects for whole Baden-Württemberg:

- Basis-DLM (without buildings) = 5.659.421 objects
- DLM50.1 = 3.872.878 objects
- DLM100.1 = 1.385.956 objects

11.8.3 *Generalisation Tool: Automatic Cartographic Generalisation*

Derivation of the Digital Topographic Maps DTK50 and DTK100 is performed by a fully automated batch process, using the Cartographic Generalisation tools of RadiusClarity and Radius ClearText from 1Spatial.

The Automatic Cartographic Generalisation tool differs from the Model Generalisation tool such that the DLM50.1 (DLM100.1) dataset is presented according the style sheet catalogue SK50 (SK100). Display of the model generalised geometries causes graphical conflicts (such as overlaps), which are detected by the

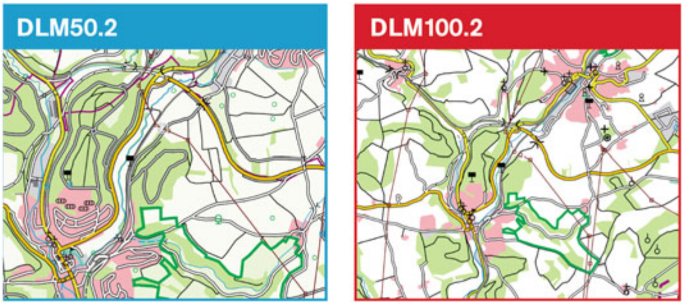

Fig. 11.20 Results of automatic cartographic generalisation

generalisation tool and solved by multiple processes. The solution to the graphical conflicts is realised by geometric and semantic generalisation rules within an adjustable workflow defining which cartographic generalisation method (displacement, simplification, typification, removing, merging, enlargement, amalgamation or enlargement) has to be used to solve a graphical conflict.

The tool uses a variety of solution methods. On the one hand there are fixed algorithms, such as the Visvalingham algorithm for simplification of areas or the Plaster algorithm to simplify lines. The AGENT-technology (Acronym for: Automatic Generalisation New Technology) has also been used. This technology allows us to model a map object as an agent which may conflict with other map objects, controlled by constraints to reach a given target. By considering these constraints, minimum size, minimum distance, minimum edge length, priority of the object, displacement and movability), the agent can search for the most optimal solution for the map object. The agent has a fixed life cycle to achieve the most suitable result to solve a graphical conflict by analysing various possible solutions (plans). Each plan is evaluated using a “happiness factor”. After a specified number of trials all results are compared to a predefined required minimum value. If the agent does not reach the desired “happiness factor”, the conflict is not resolved automatically and has to be edited manually.

Due to the variability of the generalisation tool it is possible to define each condition, parameter and workflow scale independently. So each map scale can be processed with these generalisation tools. The complete generalisation process consists of 50 individual sub-processes which can be built up one after the other. The sequence of the sub-process steps is variable and can be managed in many ways. Important generalisation steps can be repeated several times. The result of each sub-process is stored and can be used as a fall back to resume the process (Checkpoints) which allows a thorough review at the end. The processing run time of Automatic Cartographic Generalisation for the whole Baden-Württemberg is about 5 weeks for the DLM50.2 and about 3 weeks for the DLM100.2 (Fig. 11.20).

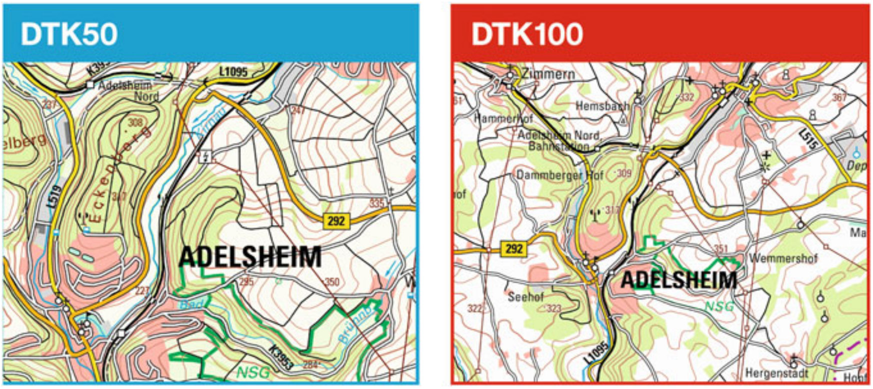

Fig. 11.21 DTK50 and DTK100 Standard Version

11.8.4 Summary

The generalisation processes for the creation of DTK100 and DTK50 (Fig. 11.21) has been used in production since summer 2011 and summer 2012 in Baden-Württemberg. Since at present not all generalisation conflicts can be solved satisfactorily in the Automatic Cartographic Generalisation process, a corresponding manual post-processing effort using the representation component is necessary. Therefore, further development and improvement of the generalisation tools are extremely important to the project. Overall, the solution already shows today that the processing and production times due to the automatic generalisation processes can be significantly reduced. So the long term aim of Baden-Württemberg is an improvement in the currency of topographic maps for all scales from 5 years to 1–2 years.

11.9 Multi-Scale Data in Spatial Data Infrastructures: Developments in INSPIRE at the JRC

Julien Gaffuri and Katalin Tóth

INSPIRE is a European directive establishing a European spatial data infrastructure (SDI). The main purpose of this infrastructure is to improve the availability and interoperability of European spatial data for environmental applications, in particular "(…) to ensure consistency between items of information which refer to the same location or between items of information which refer to the same object represented at different scales." (INSPIRE 2007, article 8.3). Different developments have been undertaken over the last few years to achieve this goal—multiple representation and generalisation have been identified as key elements for INSPIRE developments (Tóth 2007).

This section presents two fields where generalisation and multi-scale modelling were used. The first one concerns the development of interoperability target specifications, where various modelling patterns for multiple representation have been used. The second concerns a new web framework that supports the publication of vector data. For the latter we show the crucial role of generalisation and multi-scale modelling.

11.9.1 Multiple Representation in INSPIRE

11.9.1.1 The INSPIRE Data Specification Challenge

The significant achievement of INSPIRE is the release of harmonised specifications for European spatial data. These data specifications concern 34 themes such as transport networks, hydrography, geology, weather, land use, sea regions, species, etc. (INSPIRE 2007, annexes). An important challenge has been to deal with this multi-disciplinary data. For each theme, various communities with different needs and cultures have been involved. Depending on their own use cases and their specific scales of interest they often have different ways to represent the same real world objects. Overall each theme specification covers a wide thematic area and range of scales. Figure 11.22 proposes a schematic view of the INSPIRE scope and the applied scales. Cross-thematic use cases exist where two INSPIRE themes overlap—for example, the intersection between the themes transport and hydrography corresponds to water transport infrastructures.

11.9.1.2 Multiple Representation Modelling Patterns in INSPIRE

The following multiple representation modelling patterns have been used in the INSPIRE data models:

1. *Scale dependent annotations*. A commonly used modelling pattern in INSPIRE is the explicit annotation of spatial representations with their relevance to scales. For example, geographical names are characterised by a range of scales for which a given instance of a name has to be displayed. Statistical units may be represented by different geometries depending on scale extents. Some reusable scale extents are represented as lists of codes (e.g. the NUTS levels and the CityGML levels of detail). Other indirect indicators of scale are defined at the level of the datasets using specific metadata elements defined by the ISO 19000 standards series (e.g. resolution and spatial accuracy).
2. *Profiles*: When significantly diverse representations of the real world are required for various sets of use cases, different 'profiles' may be defined. When possible, the representations of the same entity in different profiles should be linked. For example, in the hydrography theme three different profiles are defined for mapping, network analysis and reporting. For the theme buildings

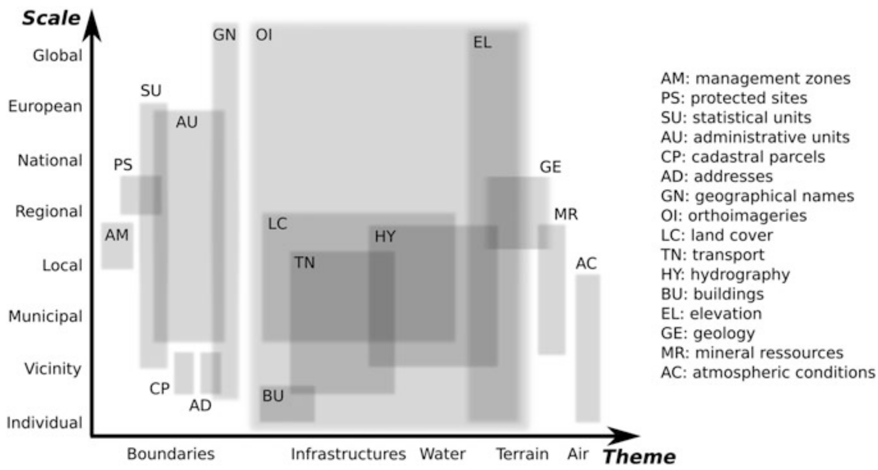

Fig. 11.22 Proposal for an INSPIRE scope representation in theme/scale space

2D and 3D data have separate profiles. Almost all data specifications are split into a simplified core profile (which contains the legally binding part of the data specification) and one or more profiles that contain additional information elements for more specific, and usually more ambitious use cases.

3. *Object and coverage views*: For many themes, depending on the scale of interest, the real world phenomena are represented as sets of objects or as coverages. For example, the themes atmospheric conditions and human health cover measurement points of physical or chemical parameters as well as interpolated coverages based on these measurements.
4. *Multi-resolution grid*: Multi-resolution geographic grids have been defined to support gridded data (orthoimagery, elevation, statistical units, etc.).
5. *Semantic level of details*: To support semantic consistency between different levels of details, hierarchical classifications are used for several themes. This includes the use of class hierarchies (specialisation/aggregation relationships) and code hierarchies (e.g. the Eurostat NACE code list on economic activities). An interesting benefit of such hierarchies is to ease the mapping of existing models with INSPIRE models: when no mapping can be found with concepts for detailed hierarchical levels, mapping are often easier to find for more generalised hierarchical levels.

11.9.1.3 Scale Gap in Spatial Data Infrastructures

As illustrated in Fig. 11.22, some parts of the scale/theme space are not yet covered by INSPIRE specifications, especially for the European scale: Some themes cover the whole scale axis, while others are limited to a very detailed range of scales. The discrepancy of scale between the data demand (pan-European applications) and offer (what is available from the national data providers) may manifest in performance issues in publishing these overly detailed datasets online. While the visualisation process of gridded coverage data widely benefits from using pyramidal data structures, similar processes for vector data are still in their infancy. For this reason a new framework has been developed, which allows an efficient usage of vector data within spatial data infrastructures.

11.9.2 A Framework for Efficient Vector Data Exchange

11.9.2.1 Next Generation SDI Based on Vector Data

Nowadays, spatial data are mainly exchanged as static images. Vector data are used only on the server side to prepare static images of the objects. Spatial data infrastructures, such as INSPIRE, are expected to support the online publication and on-the-fly utilisation of raw vector data. The reason why vector data are not used directly by the users can be explained by the performance: Vector datasets are too cumbersome to be exchanged over the network and rendered on-the-fly by the clients. In order to support the shift from static images toward dynamic vector data in SDIs, new techniques are required to ensure that only the relevant information is exchanged between the servers and the clients. For this purpose, a framework based on the following components has been developed (Gaffuri 2012):

- *Scale reference system* (*SRS*). Besides the well-known coordinate reference system, there is a need for a well-defined "scale reference system" to properly index vector data to their scale range of relevance. Taking into account the Y-axis of Fig. 11.22, twenty-one static scale levels have been defined from a local to a global scale, using a factor two between consecutive scales.
- *Generalisation methods*. In order to reduce the size of the vector data to be exchanged, model generalisation transformations are used on the server to reduce the semantic and geometric level of detail of the raw vector data. Multi-scale spatial databases are built with a suitable generalisation process for each scale level defined in the SRS.
- *Vector data tiling and spatial indexes*. In order to improve the performance of the exchange and deliver vector data only for the client's area of interest, vector tiling is used to decompose large vector objects into smaller pieces. For small and compact objects, a spatial index service is used. The tiling grid and the spatial index structure are both based on the same quad-tree structure comparable to image pyramids used in traditional raster web-mapping servers. These techniques together with generalisation allow a space/scale indexation of the data.
- *Vector and style formats*. In order to improve the vector data transfer, compact vector and style formats such as GeoJSON are used.

- *Client/server for vector data*. Specific servers and clients able to deal with these spatially indexed multi-scale vector data are used. The main capability of these components is to be able to exchange the relevant data according to the client's spatial extent and zoom level. "Scale-aware" network protocols are used. Of course, the client has the ability to render vector data on-the-fly.

11.9.2.2 Prototype and Tests

A prototype for this spatial data infrastructure based on vector data has been developed and tested. The Java4INSPIRE library was used to transform existing datasets according to the corresponding INSPIRE specifications. For model generalisation transformations the GIS Java libraries JTS and Geotools were enriched with vector tiling and spatial indexing algorithms, together with some model generalisation algorithms (object simplification, aggregation and clustering). For the GeoJSON encoding, the MapFish library was used. Two different clients were developed: one was a Java applet, and the second was based on HTML5 canvas element using the GWT library.

This prototype was successfully tested for several use cases based on Eurostat's statistical units datasets, the transport datasets of the European Environment Agency and the species distribution dataset of GBIF (http://www.gbif.org/) shown in Fig. 11.23. The raw dataset of the latter consists of 80,000 species observations represented as points. A simple clustering method based on a distance threshold is used to aggregate observations that are too close for each zoom level. The details on these observations are available only when zooming-in. In order to better visualise the distribution of these observations, a heat map visualisation method is made on-the-fly on the client side. Thanks to the generalisation and the spatial indexing, the data exchanged between the server and the client are thin; their transfer and rendering remain fast.

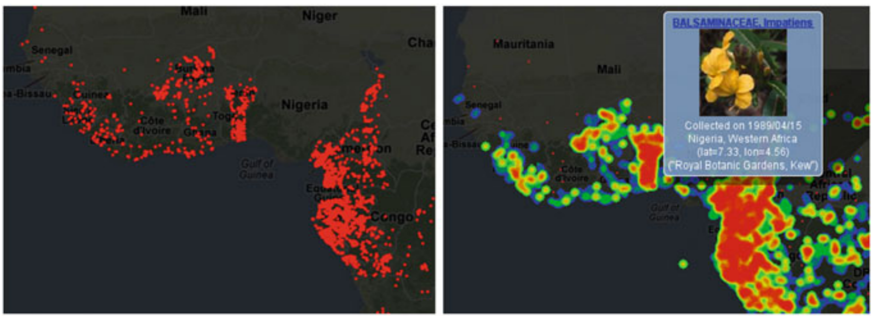

Fig. 11.23 Species distribution: raw data (*left*) and generalised data (*right*)

11.9.3 Next Steps

INSPIRE's data specification development process shows that further actions may be undertaken to progress towards a better formalisation of multiple representation modelling patterns. Standards on geographic information should better address the way to structure datasets with multiple representations. The way the level of detail of different spatial object representations is described may also be improved.

11.10 Synthesis: Recent Achievements and Future Challenges Regarding Generalisation in NMAs

This section provides a synthesis of the recent achievements, current trends, ongoing works and future challenges regarding generalisation in NMAs. This analysis is based on two sources of information. The first source of information is the set of contributions from NMAs included in this chapter—Sects. 11.2−11.7. Additional information was gathered from a larger set of NMAs during an NMA symposium held in Barcelona in March 2013, under the double cap of the ICA Commission on Generalisation and Multiple representation and the EuroSDR Commission on Data Specification, under the theme: "Designing MRDB and multi-scale DCMs: sharing experience between government mapping agencies" (ICA and EuroSDR 2013). This workshop was attended by representatives of the seven NMAs contributing to this chapter, as well as IGN-Belgium, GST-Denmark (former KMS), NLS-Finland, OSI-Ireland, IGN-Spain. Consequently, feedback from twelve NMAs is considered in total. This section is organised in two main parts followed by a conclusion. The first part summarises recent achievements and current trends regarding generalisation in NMAs, i.e. work that is already in production or has reached a certain maturity in several NMAs (11.10.1). The second part lists some challenges identified by NMAs, most of which have already begun to be tackled by work that is at a preliminary stage (11.10.2).

11.10.1 Recent Achievements and Current Trends Among NMAs

11.10.1.1 Steps in Introducing Automated Generalisation in Production Lines

Stoter (2005) and Foerster et al. (2010) identified the following steps in the introduction of automated generalisation in NMAs:

- renewing data models (from CAD-like "Map databases" to structured geographic databases, with a consistency between different levels of details),

- designing the conceptual architecture (decide what databases are derived from what databases),
- implementing generalisation processes (that actually perform automated derivation between databases), and
- managing relationships between scales.

Historically, not all NMAs have tackled these steps in the same order. The first attempts at introducing automated generalisation in production (about 10 to 15 years ago) concentrated more on the actual implementation of generalisation processes for some specific scales, widely relying on in-house developments—for example at ICC, Catalonia for 1:25,000 (Baella and Pla 2003) or at IGN, France for 1:100,000 (Jahard et al. 2003). Feedback from those first experiments (showing that automated generalisation in production was indeed possible), the availability of more and more advanced research results, and the increasing maturity of GIS software, have resulted in more NMAs taking up the challenge of automating their production lines. They have been able to do so with a much more mature and rational approach: by considering from the beginning the renewal of their data models (including consistency between levels of detail), the design of their conceptual architecture for generalisation (hereafter called "derivation scheme"), and the management of relationships between scales. The steps proposed by Stoter (2005) reflect such a rational approach to automation. Using such an approach, renewing data models was the first step. Therefore the survey by Foerster et al. (2010) logically revealed that this first step had already been completed (or mostly completed) by many of the NMAs in the survey. West-Nielsen and Meyer (2007) give a comprehensive review of how this process has been accomplished by GST (formerly KMS), Denmark. All NMAs contributing to this chapter (Sects. 11.2–11.7), or present at the Barcelona NMAs symposium in March 2013 (ICA and EuroSDR 2013), have now completed this step. However, it is interesting to note that ICC, IGN-France or OSGB have completed this step quite recently (Sects. 11.2, 11.3 and 11.5, respectively): this is because their first effort focused on the implementation of actual cartographic generalisation processes for some specific scales using in house developments, which necessitated well structured source databases, but not necessarily consistency between data models at different levels of detail.

11.10.1.2 Architectures for Generalisation and "Derivation Schemes"

The survey by Stoter (2005) was one of the first to analyse the conceptual architectures for generalisation in NMAs, which also resulted in promoting some already existing but not so widely used terms within the NMA community, which have now become widely accepted: DLM vs. DCM, model vs. cartographic generalisation, or star vs. ladder approaches.

DLM stands for "Digital Landscape Model", a term initially proposed by Grünreich (1985), which describes a geographic database where feature geometries and attributes are stored without consideration of any cartographic symbols.

This is different from DCM which stands for “Digital Cartographic Model”: a database where the geometry of the features has been distorted or displaced to take into account the scale and cartographic symbols with which they will be displayed (Weibel and Dutton 1999, and Sect. 5.6 of this book). Model generalisation occurs between one DLM and another DLM. It is driven by the data model and resolution of the source and target data. Cartographic generalisation seeks to produce a DCM from a DLM of the same resolution or from a DCM of lower resolution, while taking into account symbol widths, legibility and separation thresholds. Finally, the terms “star” and “ladder” derivation have been introduced by Eurogeographics (2005) to refer to the derivation schemes according to which the databases of different resolutions are derived from each other within a NMA. In every case a “base DLM” exists, that corresponds to the finest level of detail. In the “star” approach every DLM or DCM is derived directly from the base DLM through independent generalisation processes. In the “ladder” approach, the base DLM is generalised into a coarser DLM (resp. DCM), which is in turn generalised into a coarser DLM (resp. DCM), etc. Again, in comparison with the survey of Foerster et al. (2010), progress has been made within NMAs regarding the choice of derivation schemes, since all NMAs contributing to this chapter or present at the Barcelona workshop have now completed this step.

OSGB is the only NMA among the contributors to this chapter that chose a star derivation scheme (Sect. 11.5): all coarser DLMs are derived from the so-called “detailed” DLM, which indeed includes not only the features that are captured from the real world, but also computed enrichments that will support the generalisation processes.

All other NMAs have chosen an approach that mixes star and ladder approaches, which confirms the trend already noticed by Foerster et al. (2010). Some NMAs still maintain several DLMs of different levels of detail independently (each of them acting as a base DLM in an independent derivation scheme), and identify the merging of their derivation scheme as a necessary further step. This for example, is the case of ICC, and IGN France (Sects. 11.2 and 11.3 respectively).

11.10.1.3 Linking Databases and Managing Incremental Updates

In addition to choosing a derivation scheme, NMAs also have to decide whether they want to maintain links between their different DLMs and DCMs, thereby resulting in a multi-representation database (MRDB). Creating and maintaining links presents several advantages, summarised by Mustière and van Smaalen (2007) as being: (1) ease of database maintenance and propagation of updates, (2) assessment of quality (when existing databases are linked, knowing which is of better quality), (3) potential for more efficient applications using these databases (including more powerful analysis and visualisation tools).

The proportion of NMAs that have decided or begun to implement the MRDB approach at least for a part of their DLMs and DCMs has increased between the survey by Foerster et al. (2010) and the 2013 Barcelona NMAs workshop: eight

NMAs out of twelve in 2013, vs. five NMAs out of eleven in 2010. The links between datasets can be built in two ways. The first way is to keep derivation links during the generalisation process, when one database is derived from another. This approach has been chosen by swisstopo, OSGB and AdV for all their DLMs and DCMs (Sects. 11.4, 11.5 and 11.8 respectively), and by IGN-France for the links between its DLMs and the DCMs that are semi-automatically derived from them (Sect. 11.3). The second way to build links is to use data matching techniques between two existing datasets. This approach has for instance been chosen by ICC (Sect. 11.2). Both techniques can be combined: NLS Finland designed an automated process to derive its 1:100 k DLM from its 1:10 k DLM, but guided the automated hydrographic network selection via its existing, interactively derived 1:100 k DLM. Only segments that match a feature of the existing 1:100 k DLM are retained (Pätynen and Ristioja 2009).

Although more and more NMAs have chosen the MRDB approach, feedback from MRDB management show mixed results to date. A few NMAs have successfully implemented automated update detection between DLMs at different levels of detail (IGN Belgium, GST Denmark). Updates are introduced in the more detailed DLM and the relevance of propagating them to coarser DLMs is automatically checked according to derivation rules, before the propagation is interactively performed. Automated update propagation between a DLM and a DCM of the same level of detail has been implemented (for example at IGN-France for the 1:100 k (Sect. 11.3). But the feedback from these update propagation experiments are not without reservations due to the lack of tools that can deal with links and objects life-cycle management in existing commercial GIS software. NMAs would like to define their own unique identifiers, while most commercial GIS work with internal identifiers. Human operators capturing updates in the base DLM often prefer to delete and re-create an object that had its geometry modified instead of modifying it, which results in missing update information. This could be partially overcome by more ergonomic tools to assist updates capture, although the definition of an update as opposed to a deletion and creation is in itself quite fuzzy. For IGN Belgium, who had set up an automated updating process from DLM 1:10,000 to DCM 1:50,000, these challenges meant that they went back to an interactive propagation approach supported by a semi-automated detection (Féchir 2013). Transformations performed on the data model of the base DLM are also difficult to propagate, which also led GST Denmark to reapply a complete derivation of their base DLM to coarser DCMs in 2013, after several successful updating cycles between 2010 and 2013 (Danish Geodata Agency 2013). We also note that fully automated incremental update propagation requires the NMA to decide what specific computed enrichments should be kept in target DLMs and DCMs (e.g. partitioning information), and what information should be stored regarding the applied generalisation process, so that this process can be locally re-applied while minimising the modifications to surrounding features. To date, only partial solutions have been experimented with.

As discussed in the 2013 Barcelona NMAs workshop, the decision to apply incremental updates rather than regularly perform a complete re-derivation is dependent on several factors, among which: (1) the cost of the derivation (the more manual edits, the more costly re-derivation becomes), and (2) the willingness to provide change only updates to users so that they can in turn propagate the updates on thematic data that they have captured based on the NMA dataset. As an example, Kadaster NL chose the complete re-derivation approach for their 1:50 k DCM, since the derivation process is fully automated and they do not consider the 1:50 k DCM as a suitable backdrop dataset to which users can add their own data—although it can be used as a backdrop map (Sect. 11.7).

11.10.1.4 Implementing the Generalisation Process

Of all steps needed to introduce automated generalisation in NMAs, implementing the generalisation process for the production of topographic maps has been the most comprehensively tackled topic of NMAs and researchers. While the survey by Foerster et al. (2010) stated that "full automated generalisation processes do not exist" and that only five out of the eleven considered NMAs had made major steps towards automation, in 2013 major progress has been achieved. Eleven of the twelve NMAs attending the 2013 NMAs workshop in Barcelona have implemented automated or semi-automated generalisation processes. Fully automated generalisation processes now exist in production. This is the case in OSGB Great Britain in the derivation of a "light" 1:25 k DCM from a mixed 1:1.25–1:10 k DLM (Sect. 11.5), at IGN France for the derivation of a "light" 1:25 k DCM from a 1:10 k DLM (Sect. 11.3), and at Kadaster Netherlands (Sect. 11.7) for the derivation of the 1:50 k DCM from a 1:10 k DLM, a scale change that is traditionally considered difficult to achieve.

This full automation was achieved while accepting compromises in terms of cartographic quality and differences compared to existing manually derived products. Another currently existing approach, sometimes in the same NMAs, is to include limited manual editing in order to reach different standards of cartographic quality. One of the factors that affect the feasibility of complete automation is the kind of features present in the target DCM and their representation. To date, manual edits are still needed to reach a good cartographic quality when individual buildings are kept in dense urban areas, which mainly explains the inclusion of manual edits in the production lines of the standard 1:25 k of ICC, IGN France or swisstopo (Sects. 11.2, 11.3 and 11.4 respectively). It is also worth to notice that preserving the topological relations in the target DCM enables its further generalisation in a ladder scheme or its automated incremental updating, but potentially increases the amount of manual editing required.

Automated generalisation processes currently in use in NMAs are based on commercial GIS software distributed by 1Spatial, Axes-Systems, ESRI and the University of Hanover, that have been customised to take into account the specifics of each NMA, either by the vendors or by the NMAs themselves, relying partly on recent research results. These efforts in development, particularly intensive over recent years, have helped close the gap identified by Stoter (2005) and Foerster et al. (2010) between research and production.

11.10.1.5 Producing Maps for Delivery Over the Web

Until recently, producing paper maps was the main focus of NMAs. But in parallel most of them have developed solutions that produce raster versions of their paper maps, either sold to customers or delivered over the web through geoportals. The derivation of vector DCMs has created new opportunities. The main focus is still on cartographic products similar to traditional paper maps (e.g. traditional scales are still considered), but all NMAs understand the importance of delivering both paper and digital cartographic products. Vector web versions based on the produced DCMs are now available or under development in most NMAs, who also work on homogenising the symbolisations across scales or on delivering services that enable users to customise the display, at least by interactively selecting or unselecting layers. For a few NMAs, the main delivery channel is via the web, with printing left to the users or only purpose made, possibly outsourced to third party vendors (USGS, GST Denmark).

11.10.2 On-going Work and Future Challenges Around Generalisation in NMAs

Section 11.10.1 highlighted the fact that the introduction of automated generalisation has made significant progress over the last years. The continuously increasing availability of geographic data over the web and on mobile devices, the increasing variety of associated uses and the evolution of related economical models (Chap. 1 of this book) raise new challenges for NMAs regarding generalisation. NMAs are no longer the only producers of topographic DLMs. People can use data produced by commercial producers such as Google or from collaborative projects such as OpenStreetMap. Open Street Map data suffer from strong heterogeneities (e.g. Girres and Touya 2010), and cartographic services based on both OpenStreetMap and Google data offer limited cartographic quality (only DLMs are available, which are displayed with cartographic symbols without any cartographic generalisation). Despite this, they are sometimes more up-to-date than the data produced by NMAs and can be obtained for free. This, combined with the open data movement, is raising the expectations of users in terms of currency and low cost: they expect data to be up-to-date, and free. Producing DCMs of high cartographic quality is no longer sufficient—although cartographic quality also contributes to better decision making. Therefore NMAs have begun to adapt their economic models while minimising manual post-processing and decreasing update cycles, and to study ways of delivering new services that build upon their know-how and cartographic expertise. At the same time, the economic context is forcing administrations to decrease their number of employees. As a result, NMAs are facing the challenge of producing more, with fewer resources. From this we can distil four major challenges that have been identified and partially tackled by NMAs.

11.10.2.1 Increasing Effectiveness of DLM and DCM Production

The first challenge currently identified by NMAs is to produce faster with fewer personnel and less financial resource; one ambition being to reduce the update cycles of DLMs and DCMs. This challenge already led to greater automation in production environments, and explains the willingness expressed in all NMAs contributions to this chapter to go further with automation, whilst reducing manual intervention. Producing faster with fewer resources can be achieved at least in four ways: by improving the effectiveness of the automated generalisation processes, by adopting an incremental updating approach and therefore limiting the amount of generalisation required, by improving evaluation so that manual edits can be well targeted, and by reducing the intended level of cartographic quality (traditionally very high in most NMAs) in order to privilege fitness for use.

All four approaches can be variously combined. As stated in Sect. 11.10.1.4, one of the use cases that still requires intensive manual edits is the generalisation of 1:50 k DCMs where individual buildings are kept even in dense zones. However, recent research results show significant progress in the automated generalisation of complete topographic datasets at this scale (for example see Touya and Duchêne 2011, and Sect. 7.2 of this book). In the same way, recent research has provided good results in the automated evaluation of cases that typically require manual edits (for example Zhang 2012, and Sect. 9.7 Case study I). It is likely that further performance and quality enhancement can be expected in production in the near future.

An approach using incremental updates appears promising and has now been adopted by several NMAs which has led to a real need for new tools and methods. Tools for change detection have been successfully tested in production, and the needs for further tools and methods are well identified: further work is needed to develop tools for data management in commercial software and to set up methods for completely automated generalisation of updates with minimal modifications on surrounding objects. These improvements will take time.

It is also important to consider adjustment in the quality of outputs, especially regarding the cartographic details within DCMs, while recognising that it may be worthwhile reasonably decreasing the intended cartographic quality of a DCM compared to traditional manually produced maps, in order to increase its currency. The challenge is in deciding what is deemed to be a "reasonable decrease". Although users are not necessarily aware of it, cartographic quality enables better decision making. Seeking for maps that are cartographically perfect is counter-productive. Decreasing the intended cartographic correctness of produced maps too much would decrease their usability, and result in a vicious circle where the cartographic expertise of NMAs would be lost. To find a satisfactory compromise, better knowledge of how the users actually use the produced DCMs is necessary (as was evident from the NMAs present at the Barcelona workshop). Surveys involving user panels, such as the one reported by Kadaster NL in Sect. 11.7, are a good example of user requirements acquisition.

11.10.2.2 Building SDIs that Support the Integration of Heterogeneous Data

In addition to improving the automation of their derivation processes, NMAs are working on the rationalisation of their update collection processes. Indeed, they become more and more integrators of data collected by other providers, especially administrations. NMAs are also responding to user feedback and the activities of VGI contributors. For example the USGS has recently announced experiments to use crowd sourced POIs captured by volunteers through dedicated platforms. From the generalisation point of view, this raises the question of how to describe the level of detail of the data and how to ensure relevant interactions between integration and generalisation processes. Indeed, both processes are interrelated. Generalisation can be used in support of integration to homogenise the level of detail among multi sourced data. Integrating data is a pre-requisite to better generalisation. The same issue arises in the context of the creation of the European SDI based on data provided by European states, as required by the INSPIRE Directive. Creating this framework is in itself a real challenge. An analysis of these questions was presented in Sect. 11.9, together with some proposals and preliminary results.

11.10.2.3 New Services: Automated On-Demand Derivation of Customised Cartographic Products

With the increasing availability of geographic data, more and more users (professionals as well as the public) want to combine their own or third party data with NMA data—for visualisation and automated analysis. NMAs have begun to investigate solutions to this need. One approach is to deliver specific "light" DCMs intended to act as backdrop data on which users can capture their own data or overlay existing data, as proposed by OSGB with VectorMap District (Sect. 11.5). As a complement to this, online services can be offered to customise vector backdrop maps (conventional DCMs produced by the NMA or specific light DCMs) by selecting themes to display and adjusting symbol colours, and to overlay them with third party data (via a geoportal).

An alternate, more sophisticated solution, is to build services that enable the creation of so-called "analytical maps" integrating topographic and user data in a consistent manner. Such an approach automatically assists the user in the choice of backdrop data, integrates the user's data with the NMA data, and generalises the map whilst taking into account the user's data and their relations with topographic data, and the user's requirements. Such processes are still at the research stage (e.g. Chaps. 2, 5 and 7 of this book).

Even when no additional data are considered, the production of vector DCMs instead of just paper or raster maps opens up a lot of possibilities for NMAs in terms of on-demand customisation of their printed maps. They can enable the user to choose or build an original portrayal instead of the using the standard NMA symbols—for example portrayals with customised colours, or similar to an old map. Work in this area is on-going (e.g. Christophe 2012), and some results have already been achieved (Sect. 2.7 of this book). Their incorporation in production environments is under study at IGN-France. Such forms of customisation could also include generalisation, ideally on-the-fly, either to adapt the scale or to display using the chosen symbols.

11.10.2.4 Maintaining Composite Datasets Including User Data

In addition to integrating their own data with NMA data, users may wish to maintain such composite data over time. This use case is well known by NMAs, who all have customers who already use their DLMs for such purposes, usually by capturing their own data manually with the NMAs data acting as a reference. NMAs could support such use by offering automated integration services to enable the creation of consistent composite data from existing user data. And they could support the maintenance of such composite data by delivering updates in an incremental manner, and by offering services to semi-automatically propagate the delivered updates to the user data. This again demands better models and tools to manage incremental updates in GIS and map production software.

11.10.3 Conclusions

The utilisation of automated generalisation in NMAs has recently made significant progress. Further developments are ongoing. Longer term challenges for NMAs are currently being tackled via several research studies, although they are not solely dedicated towards NMAs. The last and probably most important challenge regarding generalisation in NMAs concerns communication. It is important to carry on encouraging communication between researchers, NMAs and GIS vendors, and increase communication with users of cartographic products and services, so that future research and development efforts keep pace with user needs and expectations. Only then can we ensure that results are converted as quickly as possible into operational tools that can eventually benefit the public at large.

Acknowledgments The authors would like to thank all participants of the NMAs symposium "Designing MRDB and multi-scale DCMs: sharing experience between government mapping agencies" held in Barcelona on 21–22 March 2013 for the rich exchanges they enabled regarding current practices of generalisation and future challenges for NMAs. They represented the NMAs contributing to Sects. 11.2–11.8 of this chapter, as well as: IGN-Belgium, GST-Denmark (former KMS), NLS-Finland, OSI-Ireland, IGN-Spain.

References

Anderson-Tarver C, Gleason M, Buttenfield BP, Stanislawski LV (2012) Automated centerline delineation to enrich the National Hydrography Dataset. In: Xiao N, Kwan M-P, Goodchild M, Shekhar S (eds) GIScience 2012. 7th international conference on geographic information science, Columbus, September 2012. Lecture Notes in Computer Science, vol 7478. Springer, Berlin, p 15

Arundel ST, Whitaker S, Phillips L, Thiem P (2010) An overview of automatic digital contour production for the USGS US Topo graphics product. Paper presented at AAG annual meeting, Washington DC, 14–18 Apr 2010

Bader M (2001) Energy minimization methods for feature displacement in map generalization. Dissertation (PhD Thesis), University of Zürich

Baella B, Pla M (2001) Map names generalization at the Institut Cartogràfic de Catalunya. In: Proceedings of the 5th ICA workshop on progress in automated map generalization, Beijing, 2001. http://generalisation.icaci.org/images/files/workshop/workshop2001/baella-pla_v0.pdf. Accessed 25 June 2013

Baella B, Pla M (2003) An example of database generalization workflow: the Topographic Database of Catalonia at 1:25,000 In: Proceedings of the 7th ICA workshop on progress in automated map generalization, Paris, 2003. http://generalisation.icaci.org/images/files/workshop/workshop2003/baella_et_al_v1.pdf. Accessed 29 May 2013

Baella B, Palomar Vázquez JM, Pardo Pascual JE, Pla M (2007) Spot heights generalization: deriving the relief of the Topographic Database of Catalonia at 1:25,000 from the master database. In: Proceedings of the 11th ICA workshop on generalisation and multiple representation, Moscow, 2007. http://generalisation.icaci.org/images/files/workshop/workshop2007/Blanca_ICAWorkshop2.pdf. Accessed 29 May 2013

Baella B, Lleopart A, Pla M (2012) ICC topographic databases: design of a MRDB for data management optimization. In: Proceedings of the 15th ICA workshop on generalisation and multiple representation, jointly organised with EuroSDR commission on data specifications, Istanbul, 2012. http://generalisation.icaci.org/images/files/workshop/submissions2012/genemr2012_session4_sub4.pdf. Accessed 29 May 2013

Balley S, Parent C, Spaccapietra S (2004) Modelling geographic data with multiple representations. Int J Geograph Inf Sci 18(4):329–354

Bard S (2004) Quality assessment of cartographic generalisation. Trans GIS 8(1):63–81

Barrault M, Regnauld N, Duchêne C, Haire K, Baeijs C, Demazeau Y, Hardy P, Mackaness W, Ruas A, Weibel R (2001) Integrating multi-agent, object-oriented, and algorithmic techniques for improved automated map generalization. In: Chinese Society of Geodesy Photogrammetry and Cartography (eds) Proceedings of the 20th international cartographic conference, Beijing, 2001

Bobzien M, Petzold I, Burghardt D (2007) Automated derivation of a 1:300,000 topographic Map from Swiss DLM VECTOR 200. In: Proceedings of 23rd international cartographic conference, Moscow, 2007

Boffet A, Coquerel C (2000) Urban classification for generalization orchestration. In: ISPRS 2000. Proceedings of the 19th congress of the international society for photogrammetry and remote sensing, Amsterdam, 2000

Brewer CA, Buttenfield BP (2010) Mastering map scale: balancing workloads using display and geometry change in multi-scale mapping. Geoinformatica 14:221–239

Brewer CA, Stanislawski LV, Buttenfield BP, Raposo P, Sparks KA, Howard MA (2013) Automated thinning of road networks and road labels for multiscale design of The National Map of the United States. Cartography and Geographic Information Systems (in press)

Buttenfield BP, Stanislawski LV, Brewer CA (2011) Adapting generalization tools to physiographic diversity for the United States National Hydrography Dataset. Cartography Geogr Inf Syst 38(3):289–301

Cecconi A, Weibel R, Barrault M (2002) Improving automated generalization for on-demand web mapping by multiscale databases. In: Proceedings international symposium on geospatial theory, processing, and applications, Ottawa

Chaudhry OZ, Mackaness WA (2005) Rural and Urban road network generalization deriving 1:250,000 from 1:1250. In: Proceedings of the 22nd international cartographic conference, a Coruña, 10–16 July 2005

Christophe S (2011) Creative colours specification based on knowledge (ColorLegend System). Cartographic Journal—ICC Paris 48(2):138–145

Christophe S (2012) Cartographic styles between traditional and original (towards a cartographic style model). In: AutoCarto'12. Proceedings of the 17th AutoCarto international symposium on automated cartography, Columbus, 16–18 Sep 2012

Danish Geodata Agency (2013) ICA Barcelona—Generalisation and MRDB. Presentation at the ICA/EuroSDR NMA Symposium, Barcelona, 2013. http://generalisation.icaci.org/images/files/workshop/symposium2013/presentations/2013_03_21_BarcelonaNMAMeeting_GST.pdf. Accessed 25 June 2013

Eurogeographics (2005) Generalisation processes: a benchmark study of the Expert group on Quality. Eurogeographics internal report, 2005

Féchir A (2013) The topo-geographic inventory. Presentation at the ICA/EuroSDR NMA Symposium, Barcelona, 2013 http://generalisation.icaci.org/images/files/workshop/symposium2013/presentations/2013_03_21_BarcelonaNMAMeeting_IGNBelgium.pdf. Accessed 25 June 2013

Foerster T, Stoter JE, Kraak M (2010) Challenges for automated generalisation at European mapping agencies: a qualitative and quantitative analysis. Cartographic J 47(1):41–54

Gaffuri J (2007) Field deformation in an agent-based generalisation model: the GAEL model. In: GIDAYS'07. Proceedings of the 5th geographic information days, Münster, 10–12 Sept 2007

Gaffuri J (2012) Toward web mapping with vector data. In: Xiao N, Kwan M-P, Goodchild M, Shekhar S (eds) GIScience 2012. 7th international conference on geographic information science, Columbus, September 2012. Lecture Notes in Computer Science, vol 7478. Springer, Berlin, p 87

Gary RH, Wilson ZD, Archuleta C-AM, Thompson FE, Vrabel J (2010) Production of a national 1:1,000,000-scale hydrography dataset for the United States—feature selection, simplification, and refinement. Scientific Investigation Report 2009-5202, U.S. Department of Interior, U.S. Geological Survey

Girres JF, Touya G (2010) Quality assessment of the french OpenStreetMap dataset. Trans GIS 14(4):435–459

Gondol L, Le Bris A, Lecordix F (2008) Cartography of high mountain areas—testing of a new digital cliff drawing method. In: Proceedings of the 6th ICA mountain cartography workshop, Lenk, 11–15 Feb 2008

Grünreich D (1985) Computer assisted generalisation: Papers CERCO-Cartography Course. Frankfurt am Main, Institut für Angewandte Geodäsie

Guislain P, Lecordix F (2011) The ever improving GIS map: evolution of cartographic representations and map production processes—Lessons learnt from the IGN topographic map production project. In: Proceedings of the 25th international cartographic conference, Paris, 3–8 July 2011

ICA and EuroSDR (2013) ICA/EuroSDR NMA Symposium 2013. http://generalisation.icaci.org/index.php/activities/12-next-events-details/48-nma-symposium-2013-general. Accessed 17 June 2013

INSPIRE (2007) 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). http://inspire.jrc.ec.europa.eu/. Accessed 29 May 2013

Jaara K, Lecordix F (2011) Extraction of cartographic contour lines using digital terrain model (DTM). Cartographic J 48(2):131–137

Jahard Y, Lemarié C, Lecordix F (2003). The implementation of new technology to automate map generalisation and incremental updating processes. In: International Cartographic Association (eds) Proceedings of the 21st international cartographic conference, Durban, 2003

Jakobsson A, Ostensen O, Lovell D, Hopfstock A, Mellum R, Kruse D, Portele C, Urbanas S, Hartnor J, Bray A, Aslesen L (2013) European location framework—One reference geo-information service for Europe. In: Proceedings of the 26th international cartographic conference, Dresden, 25–30 Aug 2013

Jenny B, Hutzler E, Hurni L (2010) Scree representation on topographic maps. The Cartographic Journal 47(2):141–149. http://www.screepainter.com. Accessed 29 May 2013

Kereneur R, Thomas J, Grémeaux N (2011) Generalisation rules for ExM medium to small scale. ESDIN Project, Deliverable D10.6. http://www.esdin.eu/project/summary-esdin-project-public-deliverables. Accessed 7 June 2013

Lecordix F, Lemarié C (2007) Managing generalisation updates in IGN map production. In: Mackaness WA, Ruas A, Sarjakoski LT (eds) The generalisation of geographic information: models and applications, Chapter 15. Elsevier, pp 285–300

Lecordix F, Plazanet C, Lagrange JP (1997) A platform for research in generalization: application to caricature. GeoInformatica 1(2):161–182

Maugeais E, Lecordix F, Halbecq X, Braun A (2011) Dérivation cartographique multi échelles de la BDTopo de l'IGN France: mise en œuvre du processus de production de la Nouvelle Carte de Base. In: Proceedings of the 25th international cartographic conference, Paris, 3–8 July 2011

Moore L (2011) US Topo—A new national map series. Direction Magazine. 16 May 2011. http://www.directionsmag.com/articles/us-topo-a-new-national-map-series/178707. Accessed 29 May 2013

Moore L (2013) US Topo—A new national map series, 2012 update. Directions Magazine, 14 Jan 2013. http://www.directionsmag.com/articles/us-topoa-new-national-map-series-2012-update/300690. Accessed 29 May 2013

Palomar-Vázquez JM, Pardo Pascual JE, Baella B, Pla M (2011): Generalizing the altimetric information of the Topographic Database of Catalonia at 1:5,000: classification and selection of break lines. In: Proceedings of the 14th ICA Workshop on Generalisation and Multiple Representation, jointly organised with ISPRS Commission II/2 Working group on Multiscale Representation of Spatial Data, Paris, France, 2011 http://generalisation.icaci.org/images/files/workshop/workshop2011/genemr2011_Palomar-Vazquez.pdf. Accessed 29 May 2013

Pätynen V, Ristioja J (2009) Vector based generalization application in the production of small scale databases of the NLS of Finland. In: Proceedings of the 24th international cartographic conference, Santiago, 15–21 Nov 2009

Petzold I, Burghardt D, Bobzien M (2006) Workflow management and generalisation services. In: Proceedings of the 10th ICA workshop on generalisation and multiple representation, Portland, 2006. http://generalisation.icaci.org/images/files/workshop/workshop2006/ICA2006-Petzold.pdf. Accessed 29 May 2013

Plessis O (2011) New cartography of France at 1:100,000 scale: an IGN mapped vector database. In: Proceedings of the 25th international cartographic conference, Paris, 3–8 July 2011

Punt E, Watkins D (2010) User-directed generalization of roads and buildings for multi-scale topography. In: Proceedings of the 13th ICA workshop on generalisation and multiple representation, Zürich. http://generalisation.icaci.org/images/files/workshop/workshop2010/genemr2010_submission_22.pdf. Accessed 29 May 2013

Regnauld N (2011) OS VectorMap District: automated generalisation, text placement and conflation in support of making pubic data public. In: Proceedings of the 25th international cartographic conference, Paris, 3–8 July 2011

Regnauld N, Plews M, Martin P (2012) An enterprise system for generalisation. In: Proceedings of the 15th ICA workshop on generalisation and multiple representation, jointly organised with EuroSDR commission on data specifications, Istanbul. http://generalisation.icaci.org/images/files/workshop/submissions2012/genemr2012_session4_sub2.pdf. Accessed 29 May 2013

Regnauld N, Lessware S, Wesson C, Martin P (2013) Deriving products from a Multi Reso-lution database using automated generalisation at ordnance survey. In: Proceedings of the 26th international cartographic conference, Dresden, 25–30 Aug 2013

Revell P, Regnauld N, Bulbrooke G (2011) OS VectorMapTM District: automated generalisation, text placement and conflation in support of making Pubic Data Public. In: Proceedings of the 25th international cartographic conference, Paris, 3–8 July 2011

Ruas A (1999) Modèle de généralisation de données géographiques à base de contraintes et d'autonomie. Dissertation (PhD), Université de Marne-la-Vallée

Ruas A (2001) Automatic generalisation project: Learning process from interactive generalisation. OEEPE Official Publication n 39

Sarjakoski T, Sester M, Sarjakoski LT, Harrie L, Hampe M, Lehto L, Koivula T (2005) Web generalisation services in GiMoDig—towards a standardised service for real-time generalisation. In: Toppen F, Painho M (eds) AGILE 2005. Proceedings of the 8th AGILE conference on geographic information science, Estoril, 26–28 May 2005

Stanislawski LV, Buttenfield BP (2011) Hydrographic generalization tailored to dry mountainous regions. Cartography Geograph Inf Sci 38(2):117–125

Stanislawski LV, Savino S (2011) Pruning of hydrographic networks: a comparison of two approaches. In: Proceedings of the 14th ICA workshop on generalisation and multiple representation, jointly organised with ISPRS Commission II/2 Working group on Multiscale Representation of Spatial Data, Paris, 2011. http://generalisation.icaci.org/images/files/workshop/workshop2011/genemr2011_Stanislawski.pdf. Accessed 29 May 2013

Stoter J (2005) Generalisation within NMA's in the 21st century. In: Proceedings of the 22nd international cartographic conference, A Coruña, 10–16 July 2005

Stoter JE, van Smaalen J, Bakker N, Hardy P (2009a) Specifying map requirements for automated generalisation of topographic data. Cartographic J 46(3):214–227

Stoter JE, Burghardt D, Duchêne C, Baella B, Bakker N, Blok C, Pla M, Regnauld N, Touya G, Schmid S (2009b) Methodology for evaluating automated map generalization in commercial software. Comput Environ Urban Syst 33(5):311–324

Stoter J, Baella B, Blok C, Burghardt D, Duchêne C, Pla M, Regnauld N, Touya G (2010) State of the art of automated generalisation in commercial software. EuroSDR official publication no. 58 http://www.eurosdr.net/projects/generalisation/eurosdr_gen_final_report_mar2010.pdf

Stoter JE, van Smaalen J, Nijhuis R, Dortland A, Bulder J, Bruns B (2012) Fully automated generalisation of topographic data in current geo-information environments. In: Zlatanova S, Ledoux H, Fendel EM, Rumor M (eds) Urban and regional data management—UDMS annual 2011. CRC Press, Leiden, pp 109–120

Stoter JE, Post M, van Altena V, Nijhuis R, Bruns B (2013) Fully automated generalisation of a 1:50 k map from 1:10 k data. Cartography and Geographic Information Science, In press. (expected to be published in January 2014)

Sugarbaker LJ, Carswell WJ (2011) The National Map: U.S. Geological Survey Fact Sheet 2011-3042. http://pubs.usgs.gov/fs/2011/3042. Accessed 29 May 2013

Swisstopo (2013) Swisstopo journey through time http://www.swisstopo.admin.ch/internet/swisstopo/en/home/current/newproducts/20130117.html. Accessed 29 May 2013

The Federal Authorities of the Swiss Confederation (2007) Geoinformation Act. http://www.admin.ch/ch/e/rs/5/510.62.en.pdf. Accessed 29 May 2013

Thom S (2005) A strategy for collapsing OS integrated transport network dual carriageways. In: Proceedings of the 9th ICA workshop on generalisation and multiple representation, La Coruña. http://generalisation.icaci.org/images/files/workshop/workshop2005/Thom.pdf. Accessed 25 June 2013

Thom S (2007) Automatic resolution of road network conflicts using displacement algorithms orchestrated by software agents. In: Proceedings of the 11th ICA workshop on generalisation and multiple representation, Moscow. http://generalisation.icaci.org/images/files/workshop/workshop2007/Thom_ICA_Workshop.pdf. Accessed 29 May 2013

Thompson RC, Richardson DE (1999) The good continuation principle of perceptual organization applied to the generalization of road networks. In: Proceedings of the 19th international cartographic conference, Ottawa, 14–21 Aug 1999

Töpfer F, Pillewizer W (1966) The principles of selection: a means of cartographic generalization. Cartographic J 3(1):10–16

Tóth K (2007) Data consistency and multiple-representation in the ESDI. In: Proceedings of the 11th ICA workshop on generalisation and multiple representation, Moscow. http://generalisation.icaci.org/images/files/workshop/workshop2007/Toth-ICAWorkshop.pdf. Accessed 29 May 2013

Trevisan J (2004) From DLM to multi representation DCM—Modelling an application on buildings. In: Proceedings of the 8th ICA workshop on generalisation and multiple representation, jointly organized with EuroSDR Commission on Data Specifications, Leicester. http://generalisation.icaci.org/images/files/workshop/workshop2004/Trevisan-v2-ICAWorkshop.pdf. Accessed 29 May 2013

US EPA and US DOI (1999) Standard for National Hydrography Dataset, United States Environmental Protection Agency, United State Department of the Interior, USGS, National Mapping Program Technical Instructions. July 1999. http://nationalmap.gov/standards/pdf/NHDH0799.PDF. Accessed 29 May 2013

USGS (1955) Map publication scales, book 1. Part B, Geological Survey topographic instructions, March 1955

USGS (1985) Part 2: Base preparation. Standards for 1:100,000-scale quadrangle maps. Jan 1985

USGS (1993) Part 3: feature specification and compilation. Standards for 1:100,000-scale quadrangle maps. Feb 1993

Wang Z, Muller JC (1998) Line generalization based on analysis of shape characteristics. Cartography Geograph Inf Syst 25(1):3–15

Weibel R, Dutton G (1999) Generalising spatial data and dealing with multiple representations. In: Longley PA, Goodchild MF, Maguire DJ, Rhind DW (eds) Geographical information systems—Principles, techniques, application and management, 2nd edn, vol 1. pp 125–155

Zhou S, Jones CB (2004) Shape-aware line generalisation with weighted effective area. In: Fisher PF (eds) SDH 2004: Developments in spatial data handling, proceedings of the 11th international symposium on spatial data handling, Leicester, Springer, pp 369–380