Multi-Scale GIS : From Generalisation to Data Integration

Andreas Illert

Bundesamt für Kartographie und Geodaesie
Richard-Strauss-Allee 11, 60598 Frankfurt, Germany
illert@ifag.de

Abstract

The paper analyses the requirements on automated generalisation when applied in the context of an existing multiple-resolution data infrastructure, where the small scale data sets should not be replaced but kept and updated incrementally. The situation is explained by example of the official German topographic-cartographic information system ATKIS.

1. Introduction

For usage with GIS systems spatial data is required at many levels of resolution. In the age of analogue maps these demands where met by a set of topographic map series from large scale to small scale. On transition to digital data the concept is continued. In many countries topographic data sets are maintained at several levels of spatial resolution with full coverage of the country at each level. While the creation of the data sets is simply carried out by capture from the maps at corresponding scale, the maintenance and update of the data poses increasing difficulties on the logistics of the national mapping agencies. Usually the data sets at different levels of resolution need to be updated manually in parallel.

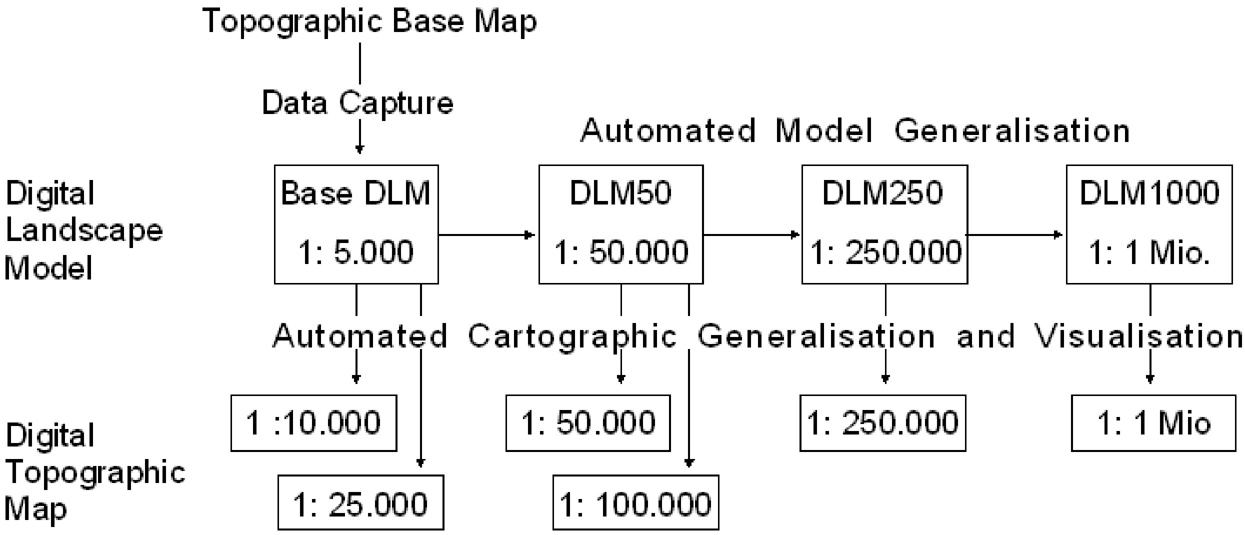
Yet the tools for automated generalisation become operational. However, the tools still require a considerable amount of interactive work. Therefore it appears not wise to abandon the existing data sets at the small scales. Instead, the generalised update information should be integrated into the existing data sets. This paper details the situation by example of ATKIS, the Authoritative Topographic-Cartographic Information System of the German mapping agencies. It does not present solutions but describes the deficits of the current implementation and points out some fields fur future research and development.

2. The initial concept

In Germany the production of official maps and data bases is shared between the agencies of the federal states and the federal government. As a rule the mapping agencies of the 16 states produce maps and data bases at scales larger then 1: 200.000, while the Federal Mapping Agency BKG takes care of the small scales. The activities of the mapping agencies are coordinated by a working committee, the AdV. This committee establishes guidelines for homogenous products and services. A prominent example for the effect of guidelines are the official topographic maps that appear with uniform layout and content though produced by several independent organisations. With the rise of GIS the AdV has set up guidelines for digital topographic data, resulting in the ATKIS concept.

ATKIS comprises of four types of digital products:

- Digital Landscape Models (DLM) describe the surface of the earth by means of structured vector data. They provide the spatial reference to many applications of the government and private users.
- Digital Topographic Maps (DTK) describe the surface of the earth in a pictorial model. The data is ready for visualisation on screen or other media.
- Digital Terrain Models (DGM) describe the third dimension of the surface of the earth.
- Digital Orthophotos (DOP) complement the ATKIS product family. The mapping agencies use the DOP internally to update the DLM. At the same time the DOP are offered to customers who are in need of original views on the landscape for their own applications.



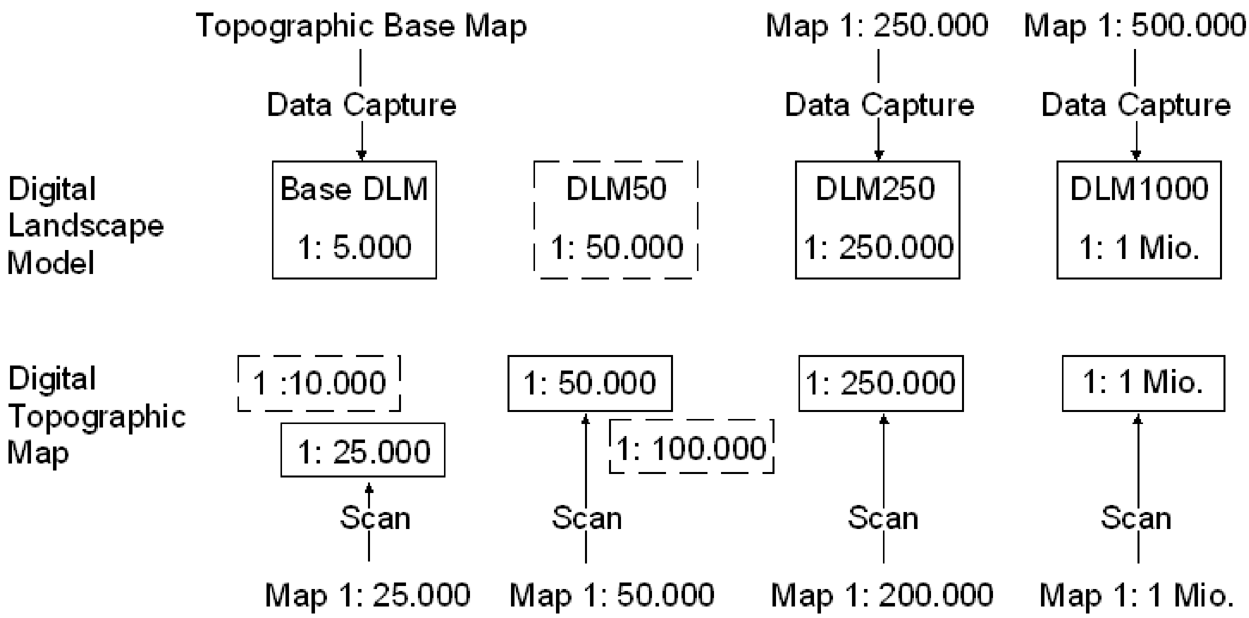
The DLM are to be provided in four levels of resolution, namely 1: 5.000 (Base DLM), 1: 50.000, 1: 50.000 and 1: 1.000.000. The ATKIS feature catalogue defines the conceptual model and the content of the DLM. The DTK are to be produced at corresponding resolutions 1: 10.000, 1: 25.000, 1: 50.000, 1: 100.000, 1: 250.000 and 1: 1.000.000. The ATKIS symbol catalogue defines the map graphics, the map layout and the portrayal rules of DTK features from DLM sources.

The initial ATKIS concept from the early 80s aimed at a continuous data flow from the base scale to the smaller scales by means of automated conceptual generalisation. For this purpose the conceptual model of the small scale DLM forms a perfect subset of the conceptual model of the Base DLM. The feature types and attributes selected from the Base DLM for inclusion to a small scale DLM keep their name, code and definition. Only the acquisition criteria and the geometry type are adapted to the lower resolution. On conversion from large scale DLM to small scale DLM the geometry is simplified, but the initial ATKIS concept aims at true position of coordinates in DLM at all levels of resolution. Unlike the DLM the DTK are subject to cartographic generalisation. The positional accuracy of the DTK is effected by processes such as displacement and exaggeration. Having the automatic procedures for model conversion and cartographic generalisation in mind the initial ATKIS concept focuses on creation and update of the Base DLM. The data sets at smaller resolution are considered to be derived from the Base DLM on demand.

3. Implementation and status

The implementation phase of ATKIS began at the end of the 80s. The data for the Base DLM is digitised from maps. The positional accuracy varies because the largest available map scale at the federal states ranges from 1: 5.000 to 1: 25.000. The GIS platform also varies between the federal states. Due to the common ATKIS guidelines the products from the different GIS software packages are as homogenous as possible, but nevertheless inconsistencies occur that are caused by specific software features such as interpolation methods or limitations on the number of vertices. The inconsistencies put considerable workload on users who are in need of full German coverage. As a consequence the German federal government has installed a geodata-centre at the BKG for centralised data collection, harmonisation and distribution of the data to the federal authorities.

At the start date of ATKIS production the tools for automated model generalisation were not considered fit for the purpose. Therefore the small scale DLM are not derived from the Base DLM but digitised from map series at corresponding scales. The geometry from the map is influenced by cartographic generalisation, hence the small scale DLM do not meet the criteria of true position. This contradicts with the initial ATKIS concept, but meanwhile it is accepted by the users as they consider the easy visualisation of data being more important at the small scales then the accurate position of coordinates.



Likewise the tools for model generalisation, the tools for automated cartographic generalisation were not considered suitable for the application at the start date of ATKIS production. In order to obtain digital topographic maps as soon as possible the printing foils of the analogue topographic maps are scanned and serve as preliminary edition of the DTK. Unless automated cartographic generalisation and symbolisation are available, the DTK have to be updated in raster format in parallel to the revision of the DLM.

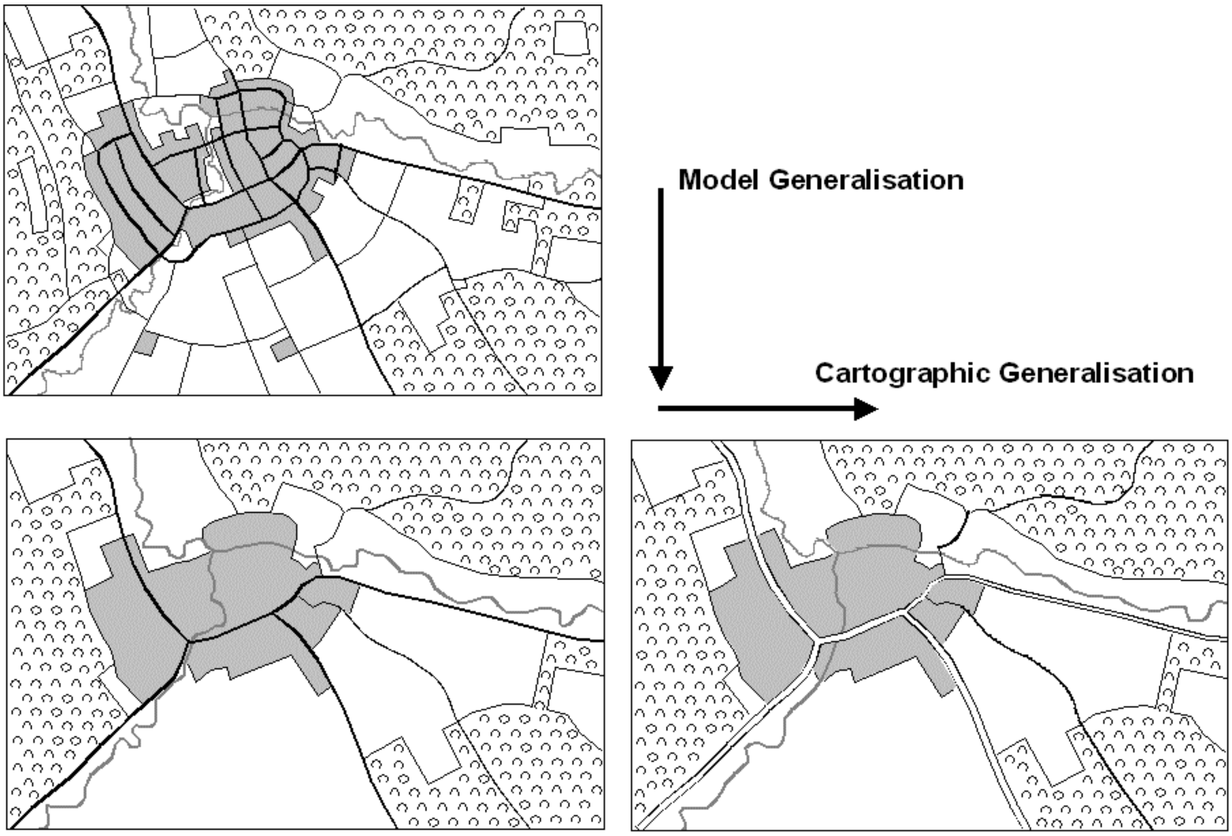
At present the Base DLM is available for the whole territory of Germany but at a reduced level of content. The state mapping agencies will extend the content to full compliance with the feature catalogue within the next years. The creation of the DLM50 has just begun, the mapping agencies still discuss on concepts and strategies. The DLM250 has been finalised a few months ago at an initial version with reduced content. The DLM1000 is available in full. Update of all DLM is in progress. The roads, railways and administrative boundaries are updated continuously, while the other features undergo a revision cycle of five years similar to the former analogue maps. The ATKIS concept itself is revised too. The experts from AdV have prepared an improved version that integrates topographic data with cadastre data and is aligned with the upcoming ISO standards.

When compared with the initial concept, a fundamental deficit of the present implementation becomes obvious. At each level of resolution within the products have been created separately from independent sources. All products comply with the ATKIS conceptual model, but at the instance level they are not linked at all. As a consequence the update process has to be performed on all products in parallel.

4. Topics for research and development

From the deficits of the current implementation two major tasks for development can be pointed out: The missing products such as the DLM50 and the DTK100 have to be created, and at the same time the missing links between the existing ATKIS products have to be introduced in order to make the update process more efficient. These two tasks have to be solved by one of – or a combination of – three steps in spatial data processing: model generalisation, cartographic generalisation and data integration.

Model generalisation creates a DLM at lower resolution from a DLM at higher resolution. In the process the amount if data is reduced but the true position of geometry is kept. Model generalisation comprises of semantic generalisation and geometric generalisation. In semantic generalisation the content of the data set and its level of detail is reduced. In geometric generalisation the spatial representation of objects is adapted to the lower resolution. Referring to the classification given in Shea, McMaster 1989 the following generalisation operators are applied: aggregation, amalgamation, selection and classification



with the semantic generalisation; simplification, smoothing and collapse with the geometric generalisation.

Cartographic generalisation creates a symbolised (graphical) representation from a DLM. The process is tied to a specific map graphic. In addition to the generalisation operators used with model generalisation the following operators are applied: merging, exaggeration, enhancement, displacement and symbolisation.

Data integration creates links between the objects of two existing data sets at the instance level. If the data sets are not already linked at the conceptual level the links between the conceptual models have to be created before. As a rule the two data sets should be at the same level of spatial and semantic resolution, otherwise model generalisation gets involved.

5. Cartographic generalisation

Of the three steps defined above cartographic generalisation can be considered the most advanced. Researchers began to develop formal approaches to cartographic generalisation already in the years preceding the introduction of GIS. The trend observed since is to divide the complex process into smaller parts where suitable algorithms can be defined and applied. Nowadays most commercial GIS software packages have implemented such algorithms and provide toolboxes for cartographic generalisation, such as the Intergraph MGE Map Generaliser or the ESRI Arc/Info tools (Lee, 1999). It is then on the user to choose a tool for a specific application, optimise the parameters and find the best strategy if two or more tools are combined. Automation of strategies constitutes a current hot topic of research in cartographic generalisation. This topic is dealt with for instance in the AGENT project (Lamy et al. 1999).

In the production of ATKIS the automated cartographic generalisation has not played a prominent role so far. Yet the most advanced application of automated cartographic generalisation is the software package CHANGE of the University of Hannover, Institute for Cartography. This software is used to generalise building features from (almost) true position at cadastre resolution to a symbolised representation in topographic maps at scale 1: 25.000. Steps in the process are selection of objects, simplification of the building polygons, amalgamation of houses and cartographic typification. Schulz (Schulz, 2000) reports on tests at the State Survey of Lower Saxony proved that the manual workload for a map sheet at scale 1: 25.000 could be significantly reduced to 8 person/hours. The software package is now implemented at several state mapping agencies.

Generalisation of buildings is only one of the many challenges within map production under ATKIS guidelines. The complex concept of ATKIS together with demand for cost reduction at the mapping agencies put a variety of requirements on generalisation tools. Such are:

a) the generalisation toolbox should provide methods for all generalisation operations.

b) automation of generalisation strategy is preferred but perhaps not necessary for certain applications. If the generalisation system is used for one single type of map the experience with CHANGE has shown that a strategy does not need to be modified after it was once optimised for a specific map graphics. However, if a generalisation system is intended for application to more than one map type, the automation can reduce workload significantly.

c) the system should be able to judge its own decisions. In order to keep manual post-processing to a minimum, the computer should make as few wrong decisions as possible, and at the same find the situations where it has not found an solution and present it to the operator for queued editing.

d) The requirements on the map graphics have to be defined by means of rules in a formal language. Such rules describe for instance the minimum dimensions of map graphics, the preservation of shape and the preservation of topologic characteristics. The rules are a prerequisite for the automation of generalisation strategies und the ability of the system to judge its results. At present the ATKIS symbol catalogues lack of suitable rules in sufficient detail.

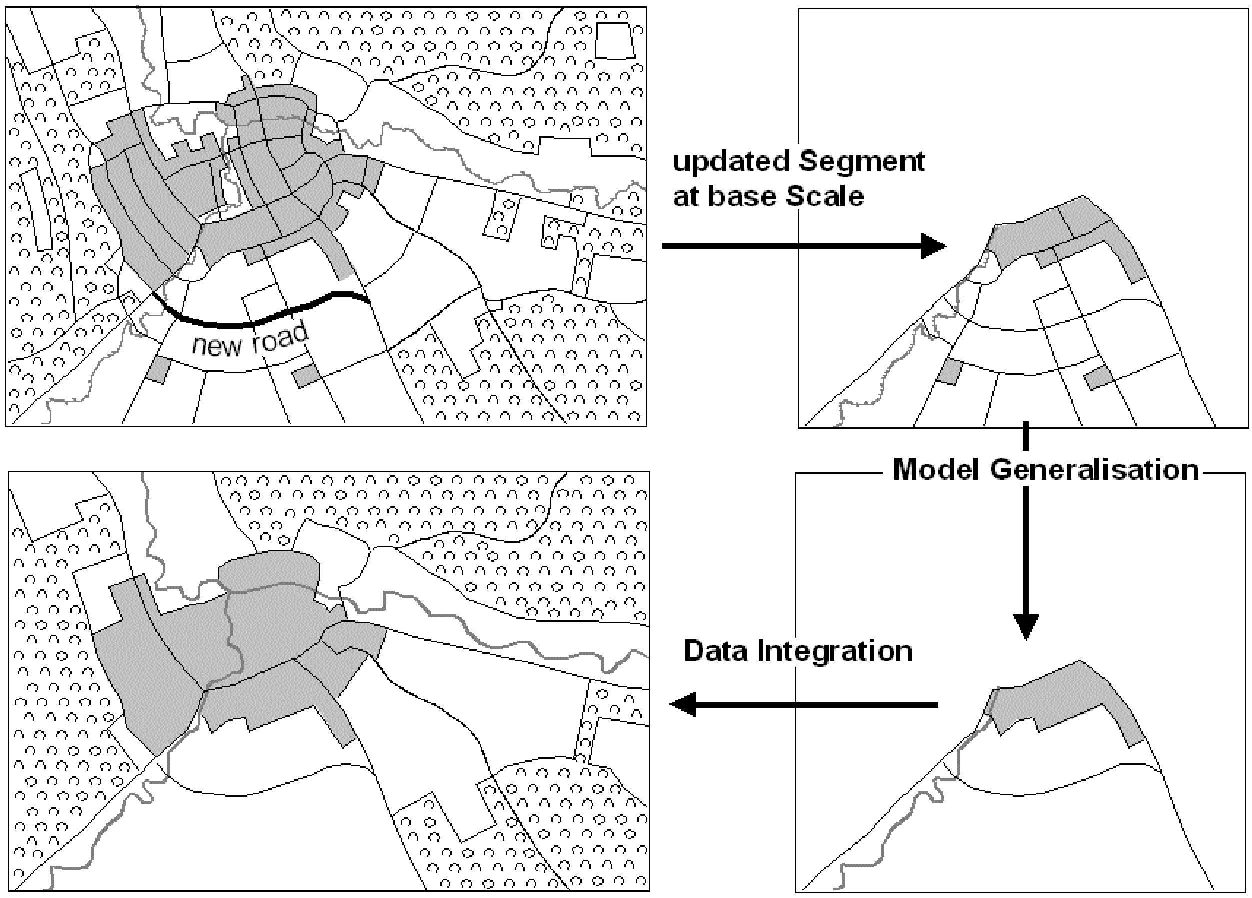
e) Future updates of the generalised data sets are hindered by the fact that the results of cartographic generalisation are in most cases not replicable. In cartographic generalisation most conflicts can be solved by more than one solution. As a consequence the automated cartographic generalisation leads to different results when, for instance, the order of objects in a data set is changed. This means that after revision of the source data set the generalisation tool will compute a data set that differs from the previous output not only at the updated patches. So, either the old data set needs to be replaced completely by the new one (with all manual post-processing to be repeated!), or the data segments with changes in the landscape have to extracted from the new generalised data set and integrated into the old generalised data set. Here, data integration will be called for.

6. Model generalisation

In principle the model generalisation can be considered a subset of cartographic generalisation as the same or at least similar tools are applied. The major difference to cartographic generalisation is that model generalisation is not constrained with map graphics. On the one hand, this makes the generalisation process easier. On the other hand the output data from model generalisation are not simple map graphics but has to be fit for usage in GIS systems. The user expects a spatial data set with reduced amount of data but similar conceptual model, data model and topology then the source data set.

The far goal of model generalisation is a series of spatial data sets at different resolutions with bidirectional connectivity between the instances. This goal complies with the concept of multiple representation topographic databases proposed by Kilpelainen 1997, Hardy 1999 and others. The bidirectional connectivity between the various object representations of the data would allow for incremental update of the generalised versions from the base data set. Having the initial ATKIS concept in mind this is exactly what the designers of ATKIS thought at.

The German mapping agencies have not used model generalisation for the production of small scale data sets yet. The one exception is the creation of a data set of administrative boundaries at resolution

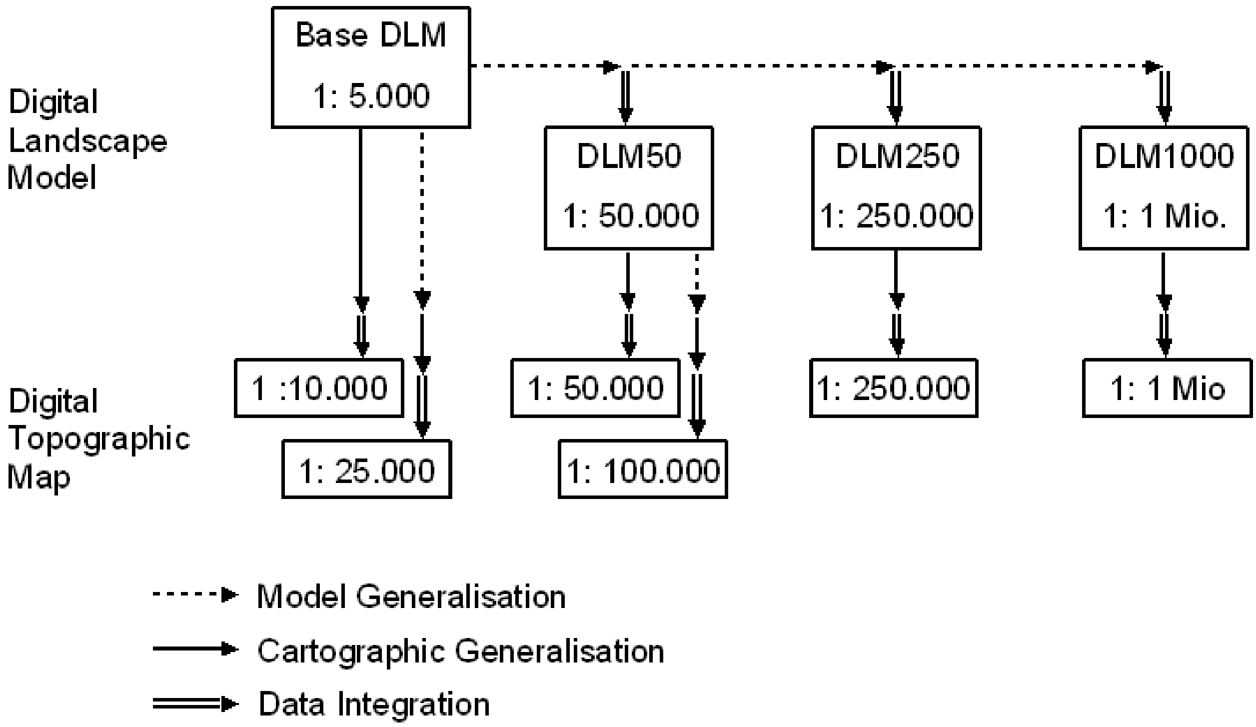


1: 1.000.000 from data at resolution 1: 50.000, but here only basic tools such as selection of polygons according to size of area and simplification of geometry using the Douglas-Peucker algorithm are applied. To overcome the deficits in model generalisation the University of Bonn, Institute for Cartography has carried out a study on automated creation of a topographic data set at resolution 1: 250.000 from the Base DLM (Morgenstern, Schurer 1999).

The study at Bonn has proved that automated model generalisation is feasible. The output still requires manual post-processing, but the degree of automation justifies the investment in hardware and software. Like cartographic generalisation the automated model generalisation may produce different results from same input if the sequence of operations or the order of data are changed, so results are not replicable. Furthermore it turned out that the rules defined in the ATKIS feature catalogue lack of sufficient detail for application of automated tools. Another problem yet to be solved is the implementation of model generalisation in a multiple representation topographic database with incremental update. Thus, in ATKIS at least two aspects of model generalisation still require further research and development:

a) the formal description of the ATKIS conceptual model needs to be refined for the purpose.
For example, a deficit of the current description is that the generalisation rules to a feature type in most cases only exploit the features' own properties such as size or dedication, but do not take into account the context. Another deficit is the lack of guidelines on how to treat the empty space after omission of area-type features.

b) in view of updating, the bi-directional connectivity between instances at the different levels of resolution has to be maintained. This connectivity information is available from the model generalisation process. It should be stored with the objects and maintained during the phase of manual post-processing.



7. Data Integration

The demand for data integration arises from the application of cartographic generalisation and model generalisation in the context of existing data sets and maps. The data producers have invested many efforts in creating these products. Yet automated generalisation will reduce the manual workload, but interactive post-processing is still required. Therefore the existing small scale data sets should not be replaced completely by new versions generalised automatically from large scale data sets with each update, but the update information is to be integrated into the existing data. One benefit is the manual post-processing being limited to the updated segments. Another benefit is that the structure of the data set is kept, allowing for incremental update service to the customers. For this purpose connectivity between the instances of the data sets at different resolutions is required. Those links may exist if the small scale data set has been created by automated model generalisation, but they are certainly missing if the data sets have been produced independently.

In the case of ATKIS, the links between data sets of different resolution do not exist. So, a first field of application for data integration is creating the links between the objects of DLM at different levels of resolution. For this purpose a combination of model generalisation and data integration has to be applied. The model generalisation is used to transform the Base DLM to a data set at same level of resolution then the existing small scale DLM. Then the data integration tools match the instances of both data sets and install the links.

Once the links are installed the propagation of update information from large scale to small scale constitutes a second field of application for data integration. The generalised update information needs to fit into the existing generalised data set. So, within a framework defined by connected instances that have not changed during update, the updated elements have to be transformed into the existing small scale DLM or map. As an option, data from the existing generalised data set might be used as a source of knowledge for the generalisation process. This however requires the generalisation operations to be modified with regard to processing of constraints.

Data integration needs to be performed at two levels, the conceptual level and the instance level. In the case of ATKIS the conceptual models of the Base DLM and the small scale DLM are already linked as they are based on the same feature definitions. Otherwise the connectivity at the conceptual level has to be created between feature types, attribute types and attribute value sets, e.g. for topographic base data of different countries or between topographic base data and road navigation data. The process of linking conceptual models is probably less suited for automation as most definitions in feature catalogues are given in natural language. As an example, the semantic integration of a data set A that contains feature types ‘river’, ‘canal (drainage)’ and ‘canal (transport)’ with a data set B containing ‘watercourse’ and ‘canal’ requires analysis of the modelling of artificial watercourses in data set B.

After connectivity is created at the conceptual level the instances can be matched. Criteria for correspondence are similarity in semantics, similarity in properties, similarity in location and similarity in topology. Note that similarity in location does not necessarily mean coincidence as displacement and exaggeration from cartographic generalisation might be involved.

Concerning the methods a knowledge-based approach should be used. In photogrammetry and remote sensing applications already exist for automated interpretation of images using procedures that match a symbolic description of an image with a landscape model derived from topographic maps. Such a software system is described for instance in (Koch, Pakzad, Tonjes 1997). At present BKG tailors this software system for application in quality assessment of ATKIS Digital Landscape Models. Having a tool for data integration in mind the system architecture and the formalised knowledge on DLM can be used from the present system while the aerial image part has to be replaced by a second data set of DLM type and the rules on matching have to be adapted to the new situation.

8. Outlook

The AdV is about to launch a call for tender for a research and development project that aims at automated creation of the DLM50 by means of model generalisation and cartographic generalisation. At the same time the ATKIS data model will be amended to allow for the storage of true position DLM geometry together with geometry generalised for maps within one data set. However, this data model of multiple geometry will not be extended to a concept that covers more than one level of resolution. As a consequence data integration will become even more important as bi-directional links between data set are not stored so they have to be recreated with each update.

References

Hardy, Paul: Active Object Techniques for Production of Multiple Map and Geodata Products from a Spatial Database. In: ICC Proceedings, Ottawa 1999

Kilpelainen, Tiina (ed.:): Cartographic generalisation in the Nordic Countries. Reports of the Finnish Geodetic Institute, No. 99:6, Kirkkonummi 1999.

Koch, H., and K. Pakzad, R. Tonjes: Knowledge Based Interpretation of Aerial Images and Maps Using a Digital Landscape Model as partial Interpretation; in: Semantic Modeling for the Acquisition of Topographic Information from Images and Maps SMATI_97, Ed.: Forstner / Plumer, Birkhauser Verlag, pages 3-19, Mai 1997

Lamy, Sylvie et. al.: The Application of Agents in Automated Cartographic generalisation. In: ICC proceedings, Ottawa 1999, pp. 1225 – 1234.

Lee, Dan: New Cartographic Generalization Tools. In: ICC Proceedings, Ottawa 1999, pp. 1235 – 1242.

Morgenstern, D., and D. Schurer, 1999: A concept for model generalisation of digital landscape models from finer to coarser resolution. In: ICC Proceedings, Ottawa, pp. 1021 - 1028

Schulz, J: Praktischer Einsatz des Gebäude-Generalisierungsprogramms CHANGE bei der Ableitung der Digitalen Topographischen Karte 1: 25.000. In.: Mitteilungen des Bundesamtes für Kartographie und Geodäsie, Band 17, Frankfurt am Main 2000, pp. 93-103.

Shea, K. S. and R. B. McMaster, 1989: Cartographic Generalisation in Digital Environment: When and How to Generalise. Proceedings of the 9th International Symposium on Computer-Assisted Cartography, Baltimore, pp. 56-67.