Karl-Heinrich Anders studied Computer Science at the University of Stuttgart and obtained the Master's degree (Dipl.-Inform.) in 1993. Until December 1999 he was staff member of the Institute for Photogrammetry of the University of Stuttgart, where he received in 2003 the PhD in Geodesy. From 2000 to 2002 he was an employee of the Z/I Imaging Corporation. Since April 2002 he is a scientific assistant at the Institute of Cartography and Geoinformatics, University of Hannover.

Dr.-Ing. Karl-Heinrich Anders

Institute of Cartography and Geoinformatics

Leibniz University of Hannover

Appelstraße 9a

30167 Hannover

e-mail: karl-heinrich.anders@ikg.uni-hannover.de

phone: +49511/762-3723

page: www.ikg.uni-hannover.de

Monika Sester, born in 1961, studied Geodesy at the Technical University of Karlsruhe and obtained the Master's degree (Dipl.-Ing.) in 1986. Until October 2000 she was a staff member of the Institute for Photogrammetry of the University of Stuttgart, Germany, where she obtained her PhD in 1995. Since November 2000 she is Professor and head of the Institute of Cartography and Geoinformatics at the University of Hanover. Her primary research interests lie in multi-scale approaches in GIS and image analysis, data generalisation and aggregation, data interpretation, and integration of data of heterogeneous origin and type.

Dr. Joachim Bobrich is geodesist and currently research associate at the federal office for cartography and geodesy (BKG) in Frankfurt/Main.

INCREMENTAL UPDATE IN AN MRDB

Karl-Heinrich Anders[1], Monika Sester[1], Joachim Bobrich[2]

[1] Institute of Cartography and Geoinformatics,

Leibniz Universität Hannover, Germany
{karl-heinrich.anders, monika.sester}@ikg.uni-hannover.de

[2] Federal Agency for Cartography and Geodesy, Germany
joachim.bobrich@bkg.bund.de

Abstract

For the generation of an MRDB two approaches can be used. The first option is the automatic derivation of all target scales from one base Digital Landscape Model (DLM). The other approach is the linking of existing DLMs and/or other data sets by a matching process. We will discuss the advantages and drawbacks of these approaches with respect to the update scenario.

1 INTRODUCTION AND RELATED WORK

A Multi-resolution/representation-database (MRDB) is a spatial database used to integrate and store data of different type, scale, thematic granularity, or accuracy which relate to the same physical phenomenon. This includes data stemming from different applications and origin, as well as data presenting and symbolizing the spatial phenomena differently. In the MRDB the connections between individual objects in the different representations are explicitly stored.

National Mapping Agencies usually provide multiple representations of their topographic data sets: on the one hand, digital landscape models (DLM) are available in different scales, on the other hand, digital cartographic models (DCM) are derived from these DLMs in order to provide a readable map. Furthermore, the DLMs can be linked with other data sets like geological maps.

If the connections between the different representations are known in terms of functional relationships – as it is the case with data of different scales – then there is the possibility to reduce the time and cost for updating the connected data sets in the future: only one landscape model (the most detailed) has to be updated manually, the derived data sets (both DLMs and DCMs) can be updated automatically by exploiting the links between the corresponding objects. The advantage of this incremental generalization approach is that an automatic update process can be implemented, which keeps the different models in a consistent state and only has to treat the affected objects. This is a strong motivation for National Mapping Agencies (NMA's) to implement an MRDB.

The concept of incremental generalisation was introduced by Kilpeläinen & Sarjakoski (1995). The idea is that an update of a change in a digital data set can be locally restricted, without having to re-generalize the whole data set. This is possible, if the generalisation process can be partitioned into modules which can be executed in a defined order. Modules can either be different thematic categories (e.g. buildings) but also spatially limited areas (e.g. mesh generated by road network, see also Galanda (2003). Harrie & Hellström (1999) present a system for the propagation of updates which is based on object specific rules. After introducing an object, a re-generalization is triggered. This approach cannot, however, guarantee, that the results of an incremental update are the same like a complete generalization of the dataset. Ellsiepen (2006) proposes a rich data structure to store the links between a generalized object and other objects together with additional information: the generalization function used, the original object(s) it was derived from, as well as the objects that have influence on it. A similar method is presented by Skogan & Skagestein (2005). Their approach is based on so-called productions that are stored in the datasets. A production is composed of an identifier of the generalization rule, that is used to create a target object, the identifiers of objects in the source dataset that are examined by this rule and the identifier of the object in the target dataset that is created. By this the history of the complete generalization is documented. Influences of updates can be found by analyzing the productions in which changed objects are involved. A disadvantage of this method is that a high amount of additional data has to be maintained. Haunert (2005) therefore proposed to exploit the information which is implicitly given with the rules: often the rule which is used for the generalization of a feature evaluates its direct neighborhood. Then, the information about the dependencies can be derived from the topology and does not need to be stored explicitly.

2 ATKIS AS AN MRDB

The ATKIS (Authoritative Topographic-Cartographic Information System of Germany) is composed of four data sets – Digital Landscape Models (DLM), which describes landscape objects in different levels of thematic and geometric detail. There is an object catalogue that describes, how the objects have to be acquired in the respective DLMs. In this way, the attributes and allowed attributes-values are fixed. Furthermore a symbol catalogue is given for each data set, which describes how to visualize the different landscape objects and to create the DCM. The different landscape models can be derived using model generalization (e.g. Podrenek, 2002, Urbanke & Dieckhoff, 2006). This ensures, that the thematic and geometric granularity of the DLMs is reduced. However, in order to adequately visualize these data sets with the defined symbology and create a DCM, they have to undergo an additional cartographic generalization process.

One option to derive the DCM is to apply cartographic generalization techniques. A second option is to use already existing cartographic presentations (e.g. digitized maps) where the cartographic conflicts already have been solved. In the first option, the links between the representations are generated by cartographic generalization procedures. In the second option the links have to be generated by matching. Both options require a high degree of automation in order to be applicable. For the updating process, both approaches have to be distinguished.

The BKG (Federal Agency for Cartography and Geodesy) is the NMA of Germany and is responsible for all authoritative cartographic products smaller than 1:100.000. The Mapping Agencies of the federal states of Germany make larger scale data sets available. All these data are compiled by the BKG to produce state-wide datasets. For sustainability and consistency reasons and to avoid multiple processing the same dataset, a MRDB is the right tool to store and keep huge datasets with different semantic, geometric and symbolic resolution up to date. So all operation can done with a high degree of automation. When changing base-data only the affected areas should be processed. Hereby time, computing power and interactive work will be saved. Hence incremental updates and propagation to more generalized levels in the MRDB are necessary. The task is to design a method that allows to limit the updates locally in order to reduce the computational and human costs.

3 ESTABLISHMENT OF LINKS IN AN MRDB

The determination of links of an MRDB depends on the underlying data sets that have to be linked. Already existing data sets can be linked using semantic and geometric matching techniques. If data sets of different thematic and geometric resolution have to be connected, the links can be generated by generalization. In this case the links imply also functional dependencies between the two data sets, that can later be used in the update process.

3.1 Links by Matching

Links can be established using matching processes. These processes have to be available for all geometric primitives, namely points, lines and polygons. There are different approaches for each cases. For all matching situations, the knowledge about the semantically comparable object classes is presumed to be known. Matching point objects can be achieved taking geometric distances into account. This approach presumes, that position and orientation of the data sets is the same. The allowed distances can vary depending on the object types (e.g. the allowed distance between matching boreholes is higher than between matching boundary stones). Other approaches try to use the local context of point objects as additional criterion (e.g. Samal et al., 2004). In this case, the requirement of known orientation can be relaxed. Matching linear objects has been investigated by many researchers. There are approaches that extend the point-based approach by using appropriate buffers. Furthermore, also geometric characteristics of the lines can be taken into account (e.g. length, orientation, number of points). In addition, also relations of the features can be used (e.g. Walter & Fritsch 1999, Mantel & Lipeck, 2004, Shi & Meng, 2006). Matching polygon objects can use the degree of overlap (e.g. measured by overlay ratios or the symmetric difference), as well as form parameters (Hild, 2003). If the matched data sets are of different scale, the matching procedures have to take these possible differences into account (see e.g. Mustière 2006). Due to typical differences in object modeling, not only 1:1-relationsships between corresponding objects occur, but more often there will be 1:n or n:m-relations.

Fig. 1, left, shows the corresponding objects of two data sets. Some of the objects have direct partners in the other data set (indicated by solid lines), there are, however, also objects which do not, and thus are linked with the superior mesh of the partner object (dotted lines).

3.2 Links by Generalization

Generating links by generalization processes has the advantage that the functional dependencies between the corresponding data sets are explicitly known and reproducible: through the process it is clear, how the objects in the target data set are represented. Generalization operations used are algorithms for model generalization (aggregation, area collapse), but also for cartographic generalization (e.g. simplification, displacement, typification). Usually, the type of generalization operation also determines the type of relationship between corresponding objects. Also in this process, all types of relations can occur, i.e. 1:1 (e.g. in case of simplification or displacement), 1:n (e.g. in case of aggregation), n:m (e.g. in case of typification).

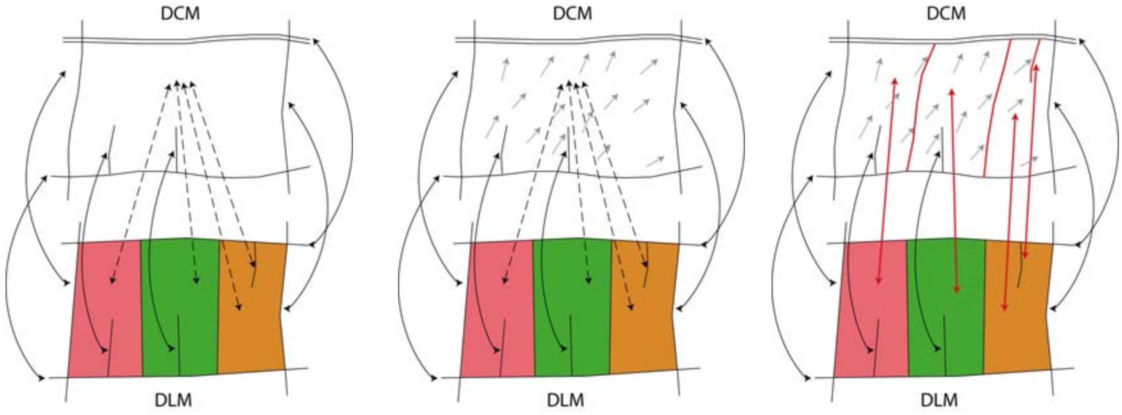

Fig. 1: left: matched Objects (solid lines), objects without direct partner in other data set are linked to mesh they are lying in (dashed lines); middle: displacement vector field generated by directly matched objects; right: propagated geometries of objects without direct partner using rubber sheeting.

4 INCREMENTAL UPDATE

The general prerequisite for an incremental update is to determine the location of the changed object and its relations to other objects, i.e. the local context that has to be taken into account for the propagation of the changes.

In order to determine the location of the update, either unique object id’s can be used, or the affected objects has to be determined by change detection with matching techniques (see e.g. Badard, 1999). In general, however, no change detection process should be needed in a consistent MRDB system because update information has to be well-defined. That means update information has to be defined by global unique object identifiers (GUID’s). In general there are three possible update types in a data base: MODIFY, INSERT or DELETE. MODIFY means an existing object defined by its GUID has to be changed (geometry or attributes). INSERT indicates that a new object with a new GUID has to be stored in the data base. DELETE means the deletion of an existing object including its GUID, so that this GUID will never be used again for other new objects. In general there is no single INSERT, DELETE or MODIFY operation in a topological consistent database. If one object is changing then all adjacent neighbor objects are changing too! E.g. the insertion of a new road brings about a modification of the superior road to which the new object is linked to; thus there is a MODIFY-operation together with the INSERT-operation.

4.1 Local context

The local context determines the objects that are affected by the change of an object. There are different approaches to determine this local context, which can be distinguished for the two MRDB-generation principles described above.

In the case of *links generated by matching*, the local context of an object correspondence can be determined indirectly – from the matching process itself. Matching objects create a displacement vector field which describes the geometric deformations an object undergoes when being transformed to the other data set. Objects that do not have a corresponding partner in the other data set can be transformed using their local context, which can be determined taking their local relationships to the matched objects into account. This can be topologic relations like containment (e.g. lake lies within forest) or connectivity (e.g. minor road joins major road). Using a topologic transformation (rubber sheeting), these relationships are preserved (see Fig. 1 middle and right).

If the links were *generated by a generalization* process, the local context of an object correspondence can be determined by inspecting all objects that are involved or affected by a generalization operation. There are different possibilities to track and control these relations (see state of the art in section 1). The generalization process has to be repeated locally for the affected objects.

4.2 Update propagation operations

4.2.1 MODIFY

In case an object is modified (see Fig. 2, left), first of all it has to be checked, if the modification only relates to semantics or also to geometry. Then, it has to be found out, if the semantic changes have influences on the geometric appearance of the objects (e.g. a change of a road type could have a change of the signature with as a consequence). If there was a change in geometry (e.g. change of the position and flow of a river), the same generalization function has to be applied as before (e.g. smoothing with Douglas-Peuker-Algorithm and given tolerance value). This generalization also has to take possibly affected neighboring objects into account. In case the links were generated by matching, rubber sheeting is applied using the displacement vector field generated by the neighboring objects. This implies, that the displacement vector field must not take the earlier relations of this object into account.

4.2.2 INSERT

If a new object is inserted (see Fig. 2, middle), it has to be assigned to the superior mesh it is lying in. This ensures, that the global context in terms of the corresponding displacement vector field is preserved. In order to transfer the new object to the other data set, the existing displacement vector field can be used. In this way the topology of the data is preserved, however, it could lead to geometrical distortions (see section 4.4). If generalization is applied, the affected objects have to be determined and used in the re-generalization process with the same parameters, taking all the objects, including the new one, into account (see Section 4.4).

4.2.3 DELETE

If an object is deleted (see Fig. 2, right), then this object will be deleted together with its corresponding object in the second data set. Also, consequences of these actions have to be taken. This could be the case when the object had formed a constraint for a related object. As an example, a road was responsible for separating two areas of similar object type. If this road is deleted, then, in turn, the areas can be merged. In the case of using the displacement vector field, it has to be checked, if the object was used to determine the vector field. If yes, then a new vector field has to be determined and the objects in the local neighborhood have to be transformed anew.

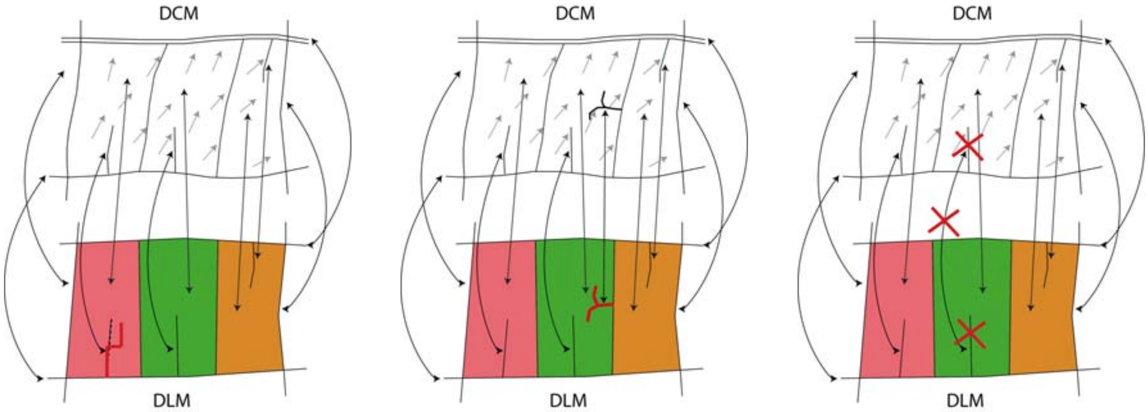

Fig. 2: Update situations: modification of red object (left), insertion of red object (middle), deletion of object (right)

4.3 Examples and discussion of rubber sheeting approach

The benefit of this approach is the fact that the relations between corresponding objects are made explicit. Based on these links, the update is relatively simple. Furthermore, the relation function (i.e. transformation between the two data sets) is bi-directional, and thus can be applied in both directions. This means that objects of data set 1 can be transferred to data set 2 and vice versa. Fig. 3 shows an example of a situation before and after application of the displacement vector field. The object is nicely adjusted in its local environment.

The drawback of using the displacement vector field is twofold: first of all, it could lead to deformations of objects which is especially severe for object with distinct shape (e.g. circular or rectangular) (see Fig. 4). Furthermore, the update can lead to the effect that spatial conflicts occur between now too close objects.

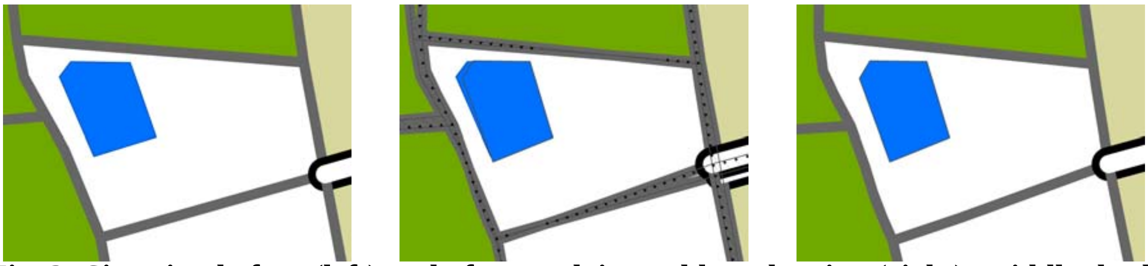

Fig. 3: Situation before (left) and after applying rubber-sheeting (right); middle: both situations and displacement vectors

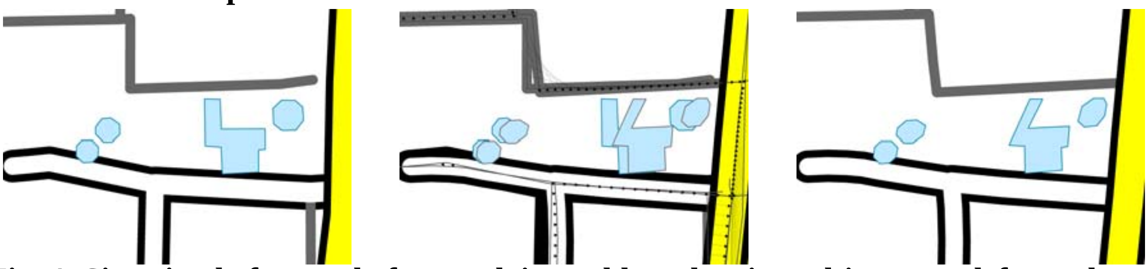

Fig. 4: Situation before and after applying rubber-sheeting: objects are deformed.

4.4 Solving conflicts with local generalization operation

In cases of local conflicts or deformation of objects, we propose to include a local generalization process instead of a rubber sheeting, namely a displacement that can take object characteristics into account. First of all, possible conflicts have to be detected. This can be achieved by checking, if constraint violations or overlaps occur after applying the rubber sheeting. In this case, all the affected object have to be selected and have to undergo a displacement operation. The determination of the affected objects is straightforward using the objects in the local neighborhood and their corresponding objects in the other data set. The objects of the local neighborhood have to "know" their generalization behavior, including their influence range. Also their neighboring objects have to be taken into account; the neighborhood has to be expanded until a "boundary" of superior objects is found, that has the effect of being able to limit the influence range of generalization process. These objects are then fed into the displacement process, in our case the program PUSH (Sester, 2005). The program is able to model the behavior of objects in terms of stiffness and positional variability. Fig. 5 visualizes that the insertion of a new road (in blue) leads to spatial conflicts with its neighboring roads, as the space is too small to adequately visualize all three roads. A displacement with the given vector field in this case can not solve this conflict. Therefore, in that local context area the objects undergo a displacement operation, which leads to the desired result of representing all objects adequately.

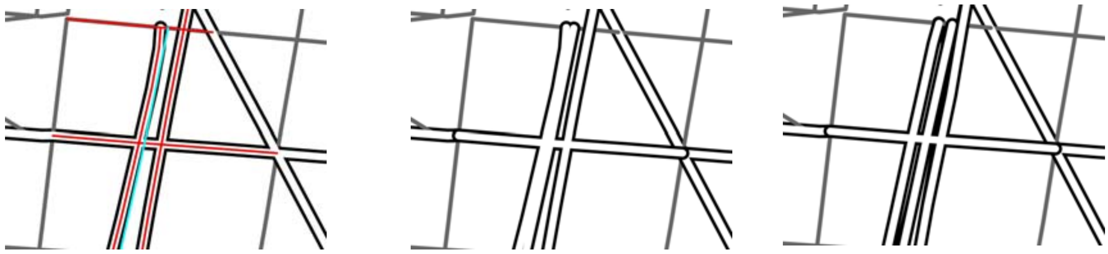

Fig. 5: Insertion of new road (blue) (left); conflicts after applying signature (middle); situation after conflict resolution with displacement (right)

5 CONCLUSION

The paper discussed the problem area of incremental updates in an MRDB. From the standpoint of a National Mapping agency there are different demands and wishes with respect to such a data structure. A NMA has to provide data sets of different scales in a consistent manner. Furthermore, there are new demands concerning the seamless integration of also other data sets into their data sets. This also includes the option of being able to transfer updates to the related data sets – but also to include updates from those data sets as well. The concept presented in this paper relies on a combination of determining links between corresponding data sets using matching techniques and local generalization operations. In this way, simple changes in the data sets can be propagated using the displacement vector field generated by matched features. In case of spatial conflicts, there is the possibility to apply a displacement operation in a local context area.

Further investigations have to concentrate on the design of the overall control strategy for this process. Also, the model of the data structure for the displacement vector field will be determined. Finally also the decision on the model of the context objects will have to be taken.

6 ACKNOWLEDGEMENTS

The research presented in this paper is funded by the Federal Agency of Cartography and Geodesy in Germany. The support is gratefully acknowledged.

7 REFERENCES

Badard, T.: On the automatic retrieval of updates in geographic databases based on geographic data matching tools. In: Proc. of the 19 th Int. Cartographic Conference, Ottawa, ICA/ACI (Eds.), 1999, pp.47 – 56.

Ellsiepen, M, 2006: Strategies for automatic and interactive generalization of topographic vector data over several scales (in German), in: DGfK-Workshop "Aktuelle automationsgestützte Verfahren zur Generalisierung topographischer Daten" (in German), Bonn, 22-23.11.2006.

Galanda, M., 2003: Automated Polygon Generalization in a Multi Agent System, PhD Thesis, Department of Geography, University of Zurich, Switzerland.

Harrie, L. & Hellström, A.-K., 1999: A Prototype System for Propagating Updates between Cartographic Data Sets. – The Cartographic Journal, 36(2), pp. 133-140.

Haunert, J.-H.: Link based Conflation of Geographic Datasets. ICA Workshop on Generalisation and Multiple Representation, 7.-8.July 2005, La Coruña/Spain.

Hild, H. [2003], Automatische Georeferenzierung von Fernerkundungsdaten. Deutsche Geodätische Kommission, Reihe C, Nr. 562, München 2003.

Kilpeläinen, T. & Sarjakoski, T., 1995: Incremental generalization for multiple representations of geographic objects. In: Muller, J. C., Lagrange, J. P. & Weibel, R. (editors) GIS and Generalization: Methodology and Practise. Taylor & Francis, London, pp. 209-218.

Mantel, D. & Lipeck, U. W., 2004: Matching Cartographic Objects in Spatial Databases, in: Int.Archives of Photogrammetry, Remote Sensing and Spatial Inf.Sciences, Vol. 35, Part B4, pp. 172-176, ISPRS, Istanbul, 2004.

Mustière, S., 2006. Results of experiments on automated matching of networks at different scales. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 2/W40), pp. 92-100.

Podrenek, M., 2002: Aufbau des DLM50 aus dem Basis-DLM und Ableitung der DTK50 – Lösungsansatz in Niedersachsen. In: KS, Band 6, Kartographie als Baustein moderner Kommunikation, S.126-130, Bonn.

Sester, M., 2005: Optimizing Approaches for Generalization and Data Abstraction, in: International Journal of Geographic Information Science, Vol 19, Nr. 8-9, pp. 871-897, 2005.

Samal, A., Seth, S. & Cueto, K., 2004: A feature-based approach to conflation of geospatial sources. – International Journal of Geographic Information Science 18 (5): 459–490.

Shi W. & Meng, L. (2006): Some ideas for integrating multidisciplinary spatial data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 2/W40), p. 30-35

Urbanke, S. & K. Dieckhoff, 2006: Das AdV-Projekt ATKIS-Generalisierung, Teilprojekt Modellgenerali-sierung, Kartographische Nachrichten 4, 2006, S. 191-196.

Walter, V. & Fritsch, D., 1999: Matching spatial data sets: a statistical approach. – International Journal of Geographical Information Science 13(5), pp. 445-473.