Towards extraction of constraints for integrating environmental spatial data in digital landscape models of lower resolution – a work in progress

Christian Stern[1], Monika Sester[2]
[1]Faculty of Geomatics, Karlsruhe University of Applied Sciences, 76133 KARLSRUHE, Germany
Christian.Stern@hs-karlsruhe.de
[2]Institute of Cartography and Geoinformatics (ikg), Leibniz Universität, 30167 HANNOVER, Germany
Monika.Sester@ikg.uni-hannover.de

INTRODUCTION AND RELATED WORK

Nowadays an abundance of environmental spatial data like nature sanctuaries, landscape conversation areas and water protection areas, is publicly available. Mainly online map services have been set up by environmental authorities to fulfill the requirements set by several laws like the Directive to establish an Infrastructure for Spatial Information in the European Community (INSPIRE), published in May 2007, or its adaption for the German national legislation, the "Geodatenzugangsgesetz" (GeoZG) released in February 2009 (INSPIRE, 2007; GeoZG, 2009).

These services usually use spatial environmental data collected in high precision and overlay it on topographic data, no matter if the data is displayed in large or small scales. Figure 1 shows a cutout of such a map service published by the German Federal Agency for Nature Conservation in early 2011 at different scales. The excerpt shows the region around the city of Hannover in Lower Saxony, Germany. Nature protection areas are shown as green areas or marked as green triangles if their size is too small to present it properly. In the left part of the image there are three protection areas encircled in the vicinity of the town of Laatzen which all cross the motorway completely (scale of screenshot: 1:500.000). In the middle part of the image we see the same nature sanctuaries (at a screenshot scale of 1:200.000), giving us a slightly different impression. The areas are now rather intersecting the motorways in parts than crossing them totally. Finally, at a screenshot scale of 1:50.000 we learn that our assumption still has been wrong. The protection area marked with a dashed line is only touching the motorway.

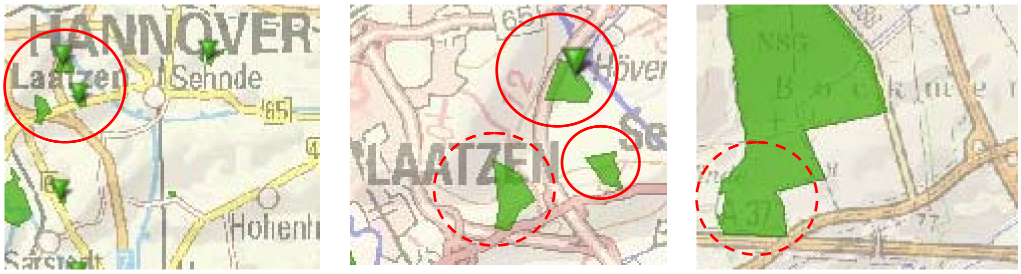

Figure 1: Nature protection areas presented at an online map service at the scales of 1:500.000 (left), 1:200.000 (middle) and 1:50.000 (right). (BfN, 2011). Note the topological relationships in respect to the motorways (encircled in red)!

The example described above clearly shows that it is not a good solution to overlay topographic raster maps of various scales with highly accurate environmental spatial data. It may lead to wrong topological relationships that do not conform to reality.

To overcome this deficit of current online map services or as a basis for printed maps in small scales that are still needed nowadays, it is necessary to integrate and generalize environmental spatial data into digital landscape models (DLM) of lower resolutions. Based on the newly derived data, new map services without the known limitations of raster maps could be created or printed maps can be compiled on demand. The goal of our research is to automatically identify constraints between environmental information and topographic data and preserve them in the integration and generalization process.

Stern (2009) proposed a concept to perform this task by using the advantages of a Multi-Agent-System. In this case, each protection area would be modeled as a so called “agent” that is aware of its characteristics and its environment and knows how to generalize itself to fit into the new landscape model of lower resolution. To solve conflicts with other entities of the system, agents have the possibility to communicate with each other. In competition or joint actions they try to achieve their goals and work out the optimal solution. Multi-Agent-Systems have already been successfully used in generalization and related processes before. For example, Galanda (2003) used such a system for automated polygon generalization, or Ruas & Duchêne (2007) reported on a prototype generalization system based on the multi-agent system paradigm, while Stern (2010) presented a solution to cluster environmental spatial data by means of a Multi-Agent-System as a preprocessing step for data integration and generalization.

There have been several proposals for using constraints in polygon generalization, but to our knowledge, nobody has tried to find them for the specific purpose to integrate environmental spatial data in digital landscape models of lower resolution yet. Using constraints for data integration has been proposed by several authors. Oosterom (2006) and Steininger & Weibel (2007) propose a taxonomy of constraints mainly targeting at integration. Hespanha et al. (2008) use OCL to formalize the constraints. Werder (2009) gives a comprehensive overview on research in constraints and the formalization of them.

Nevertheless, all these approaches have in common that we need solid constraints from which we can derive rules for the generalization task. In the next paragraphs, we will look upon how such constraints can be found to tackle the problem to integrate spatial environmental data in landscape models of lower resolution and to generalize them so that they fit to their target scale.

In our investigations, we mainly concentrate on the topological constraint “touch”, as we are interested in preserving and enforcing adjacency constraints in the process of integration and generalization. Future work will also take additional constraints (such as overlap and contains) into account.

APPROACH FOR CONSTRAINT EXTRACTION

In order to determine which topographic information typically borders environmental data, we investigate existing data sets. In the state of Baden-Württemberg, Germany, environmental data in form of protected areas is captured based on the borderlines of land parcels (LUBW, 2006). Thus, we should be able to derive rules for integration and generalization by finding out which landuse type is adjacent to the protection areas. The more often one specific landuse class is found, the more it can be counted as an important constraint for integration and generalization when topological relationships have to be preserved. Of main relevance are changes in the landuse type, because only they can be recognized in digital landscape models of lower scales.

As study area, we chose “Emmendingen”, an administrative district (German: “Landkreis”) in southern Germany, in the state of Baden-Württemberg (see figure 2). It stretches from the densely inhabited region of the Rhine valley to the sparsely populated regions of the Black Forest with its mountains and valleys covered by vast forest areas and rural landuse. This ensures that we can cover a variety of landuse types and morphological landscape patterns.

Figure 2: Location of the study area "Emmendingen" (outlined in red) in the federal state of Baden-Württemberg in southern Germany.

Among the many types of spatial environmental data we chose water protection areas as research objects. They seem to be the most interesting research object as they can differ largely in size and shape and there is no fixed rule how or by which landscape objects they have to be delimited. Seemingly, nearly any landscape object may appear as border. So the focus of this work is to determine frequencies of landuse type combinations that can be found at borders of water protection areas. This will help us to understand the nature of these areas and to derive rules for later integration and generalization processes.

As outlined before, at least in theory the boundaries of protected areas match the borderlines of land parcels. But in practice, only a small number of direct matches can be found when applying topological queries. Therefore, we extract all landuse parcels that intersect with a two-sided buffer around the outline of the water protection areas in a first approach. Following that, the outlines of the protected areas and of the landuse parcels are split into their segments. In cases where we find start or end points of our landuse parcels segments but not on the protection area segments (figure 3, left), we split the latter at that location to make sure we can find the segment in the following selection operation (figure 3, middle). To avoid getting misleading results using the "identity" operation of ArcGIS, we select all water protection area segments that lie completely on the segments of our landuse parcels (see figure 3, right). Finally, we use the identity operation that computes a geometric intersection of the water protection areas' segments and the landuse parcels joining also their attributes and keeping their relationships (see ESRI, 2011). As first result, we get the information for each segment what type of landuse we find to its left or right side.

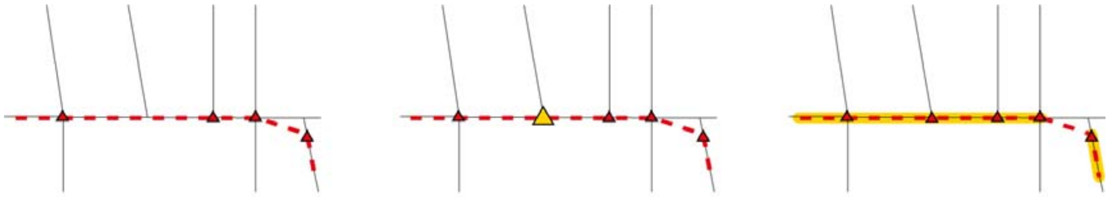

Figure 3: Process of preparing the water protection area segments for the "identity" operation. Left: water protection area (WPA) segments (red) and landuse parcel segments (black) do not match completely. Middle: the WPA's segment is split. Right: WPA segments that lie completely on landuse parcel segments are selected for the following "identity" operation.

Now we can determine the frequency of the landuse type combinations we find on both sides of our water protection area segments and sum up the segments' lengths respectively. To reduce the abundance of possible combinations of landuse types (there are 47 landuse types according to LVA-BW (2002)), we use the coarser classification of 15 "landuse classes" provided by the land surveying agency of the federal state of Baden-Württemberg (see table 1).

#	landuse class	#	landuse class
1	buildings and grounds	9	forest, wood
2	commercial and industrial areas	10	heathland, moor, swamp, fallow land
3	recreation areas or green space	11	water
4	transportation areas	12	training area
5	railway areas	13	protected area
6	farmland, vineyard	14	historical site
7	grassland	15	badlands
8	garden land		

Table 1: Landuse classes used for reduction of landuse type combinations (translated from LUBW, 2011).

RESULTS

As result we receive a matrix of 120 possible combinations (combinations of the type "1-2" and "2-1" are seen identical). Table 2 shows the resulting matrix. 65 of these possible combinations have been found in our study area (those with landuse class "unknown" because of neighboring landuse parcels lying outside of the study area excluded). If we extract the 10 most important combinations (highlighted in table 2), we still can cover 77 percent of the outlines. For the study area of "Emmendingen" we can state that 41 percent of the water protection areas' borders are a combination of transportation areas (streets, places, paths, etc.) with another class. The most important one is the combination of transportation areas with farmlands or vineyards (about 21 percent), directly followed by the landuse combination forest / forest (about 20 percent). In an additional analysis we checked the touch-relation with administrative boundaries. It turned out that they can be found in 14 percent of all protection area outlines (again excluding the landuse class combinations with "unkown" and the coincident borders at the edge of our study area). If boundaries of water protection areas follow administrative borderlines, this is most likely to be in forests (about 69 percent of these cases) or at the edge of forests (combinations of a landuse class with "forest", approx. 20 percent).

An overview on our results gives figure 4 where the detected (red) and undetected (gray) water protection area segments have been overlaid on the most important landuse classes. We term those segments "undetected", which do not have a boundary in common with a landuse class.

CONCLUSION

We have shown a solution on how to tackle the problem of extracting constraints specific to the task of integrating environmental spatial data into digital landscape models of lower resolution. The results will be used in further processing steps to control the matching process of DLM objects and protection areas: if several DLM objects are candidates for matching, then our results could be used to determine the most feasible matching partner. In the process of generalization, this matrix can help to decide which generalization operator to choose. The decision process is indicated in figure 5.

landuse class	buildings and grounds	commercial and industrial grounds	recreation grounds or green space	transportation grounds	railway grounds	farmland, vineyard	grassland	garden land	forest, wood	heathland, moor, swamp, fallow land	water	training area	protected area	historical site	badlands
buildings and grounds	0,75	0,06	0,01	3,80	0,02	0,40	0,30	0,04	0,09	0,01	0,17	-	-	-	0,46
commercial and industrial grounds	0,06	0,07	0,05	0,59	-	0,1	0,06	0,02	0,05	-	0,15	-	-	-	0,16
recreation grounds or green space	0,01	0,05	0,26	0,31	0,29	0,02	-	0,05	0,3	-	0,26	-	-	-	0,18
transportation grounds	3,80	0,59	0,31	3,38	0,73	20,70	4,25	1,41	5,14	0,00	0,4	-	-	-	3,89
railway grounds	0,02	-	0,29	0,73	0,02	0,36	0,16	-	-	-	-	-	-	-	0,19
farmland, vineyard	0,40	0,1	0,02	20,70	0,36	8,41	1,43	0,51	2,66	-	2,05	-	-	-	1,53
grassland	0,30	0,06	-	4,25	0,16	1,43	3,93	0,01	3,49	-	2,05	-	-	-	1,44
garden land	0,04	0,02	0,05	1,41	-	0,51	0,01	0,37	0,01	-	0,01	-	-	-	0,06
forest, wood	0,09	0,05	0,3	5,14	-	2,66	3,49	0,01	20,04	-	0,44	-	-	-	0,86
heathland, moor, swamp, fallow land	0,01	-	-	0,00	-	-	-	-	-	0,01	0,01	-	-	-	-
water	0,17	0,15	0,26	0,4	-	2,05	2,05	0,01	0,44	0,01	0,49	-	-	-	0,13
training area	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
protected area	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
historical site	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
badlands	0,46	0,16	0,18	3,89	0,19	1,53	1,44	0,06	0,86	-	0,13	-	-	-	0,44

Table 2: Matrix of combinations of landuse classes to be found on either side of water protection areas in the study area “Emmendingen”. The figures show the percentage of water protection areas’ border length one landuse class combination covers compared to the total length of borders. The 10 most important combinations have been shaded in blue. Only combinations of landuse changes are useful for the following steps to determine feasible partners in the matching process.

We do know the limitations of the approach presented: We only had a look at those parts of water protection area borders which are coincident with that of landuse parcels. Due to digitizing errors, we find that in many cases these lines do not fit together perfectly. Working with growing buffers, we found increasingly more water protection area border segments that can be assumed to lie on landuse parcel borders. The chart shown in figure 6 reveals that with increasing buffer distances (x-axis), we can catch more border segments (percentage of length of found segments to total protection area boundary length shown on the y-axis). In further steps, automated matching processes like those that are used for matching street networks could be adapted to this problem.

To determine the correct underlying landuse combinations for these additional border segments and to sort out the wrongly detected ones is ongoing work. Additionally, we will have to compare the results derived from our study area (Emmendingen) to those of other areas and examine if general rules can be found and transferred. Finally, we will have to research if the rules we found also apply to other environmental spatial data than water protection areas. These are the problems we will have to cope with next.

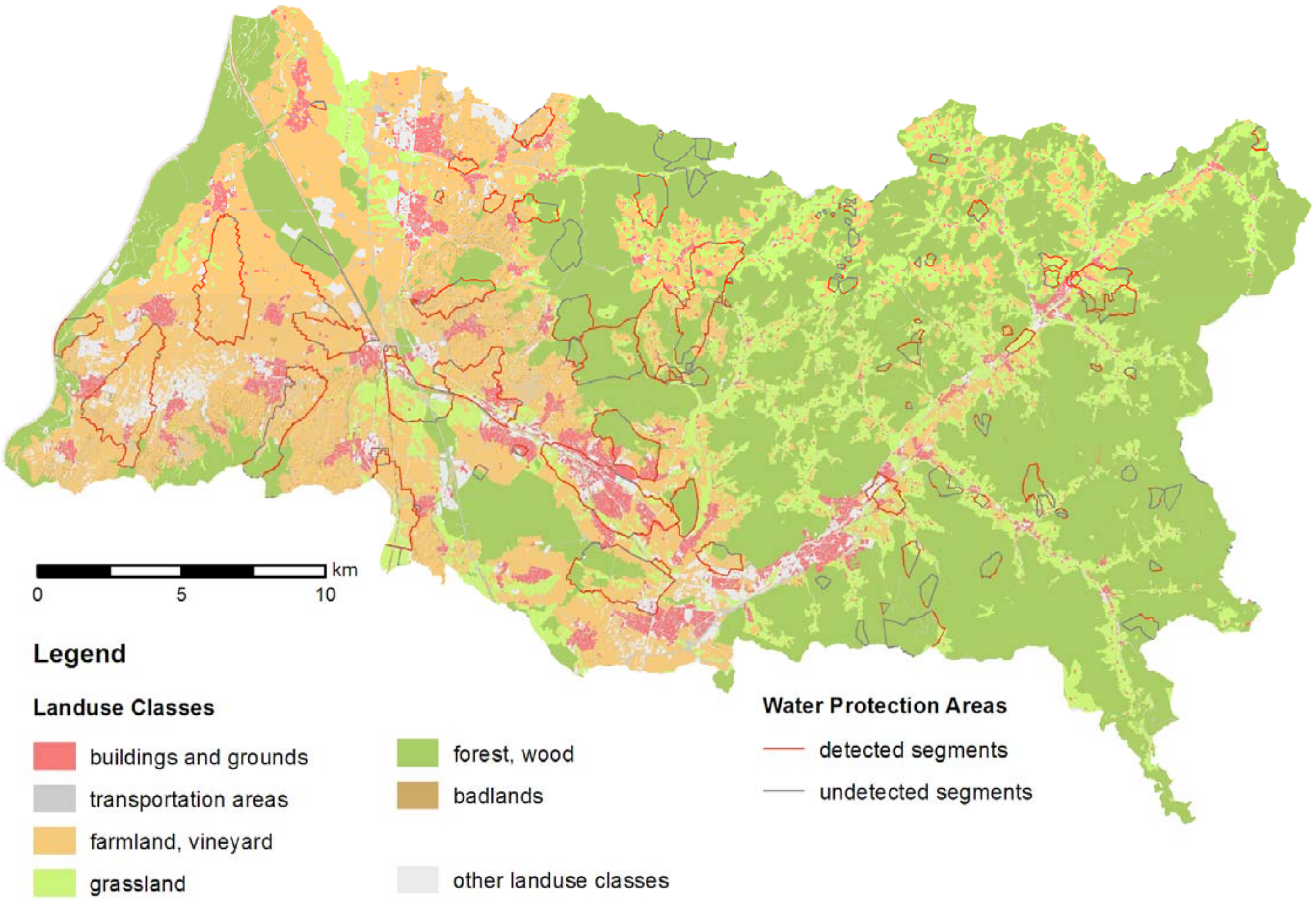

Figure 4: Analysis results in the study area of “Emmendingen”. Detected water protection area segments are shown in red, while undetected ones are displayed in gray colors.

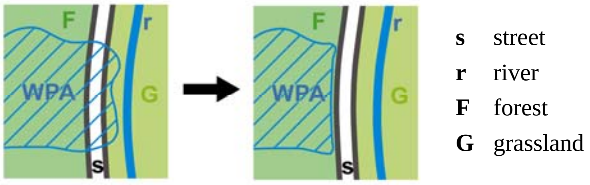

Figure 5: Decision process of matching the water protection area’s (WPA) border to the DLM object ‘street’ (s) or ‘river’ (r): From our results we know that combination sF (5,14%) is the most feasible one, compared to sG (4,25%), rG (2,05%) or rF (0,44%). Therefore, the border is matched to the street and generalized using the operator ‘enhancement’.

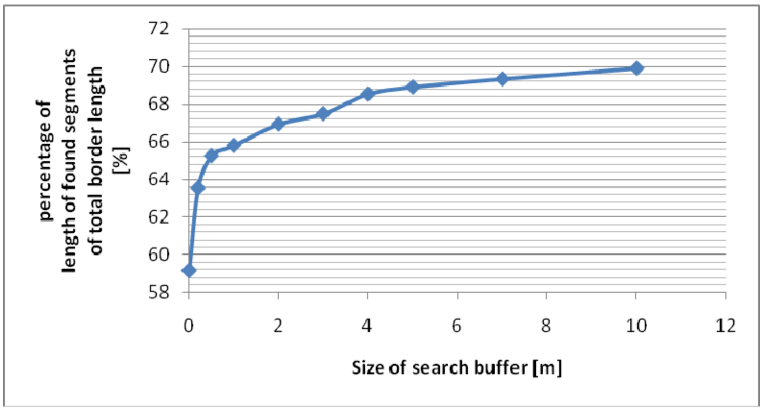

Figure 6: By increasing the search tolerance (buffer size) a higher percentage of borders can be found where we may assume that they are lying on landuse parcel borders.

BIBLIOGRAPHY

Bundesamt für Naturschutz (BfN) (ed.), 2011 Schutzgebiete in Deutschland. URL: http://geodienste.bfn.de/schutzgebiete/. (Accessed: 02-22-2011).

ESRI, 2011 ArcGIS 10 Desktop Help. URL: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html. (Accessed: 02-24-2011)

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

GeoZG, 2009 Gesetz über den Zugang zu digitalen Geodaten (Geodatenzugangsgesetz). In BGBl. I p. 278.

Hespanha, J., J. van Bennekom-Minnema, P. van Oosterom, C. Lemmen, 2008, The Model Driven Architecture Approach Applied to the Land Administration Domain Model Version 1.1 - with Focus on Constraints Specified in the Object Constraint Language, FIG Working Week 2008, Integrating Generations, June 2008, Stockholm, 19 p.

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) 14.03.2007. In Office Journal of the European Union of 25 April 2007, L 108/1.

Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW) (ed.), 2011 Metadaten. URL: http://ripsmdk.lubw.baden-wuerttemberg.de/technik/alk/nutzung.htm. (Accessed: 02-24-2011).

Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW) (ed.), 2006 Richtlinien zur Fachdatenführung im Naturschutz. Entwurf 2.2. Bearbeitungsstand vom 18.07.2006. Vorgelegt von LUBW Referat 25.2 / Referat 53.2 (ITZ). (Unpublished).

Landesvermessungsamt Baden-Württemberg (LVA-BW) (ed.), 2002 BGRUND-Schnittstelle. Schnittstelle für den Austausch von Punkt- und Grundrissdaten der ALK-BW. URL: http://www.lv-bw.de/lvshop2/ProduktInfo/testdatencd_rom/rubrik/beschreibung/format/BGRUND/beschr_BGRUND.pdf. (Accessed: 02-27-2011).

Oosterom, P. van, 2006 Constraints in spatial data models, in a dynamic context. In Drummond J., Billen R., Joao E., Forrest D. (eds.) Dynamic and mobile GIS: investigating changes in space and time. Taylor & Francis, London, pp. 104-137.

Ruas, A., Duchêne, C., 2007 A Prototype Generalisation System Based on the Multi-Agent System Paradigm. In W. A. Mackaness, A. Ruas & L. T. Sarjakoski (eds.). Generalisation of Geographic Information: Cartographic Modelling and Applications. Elsevier: pp. 269-284, 2007.

Steiniger, S., Weibel, R. 2007 Relations among map objects in cartographic generalization. Cartography and Geographic Information Science (CaGIS) 34 (3): pp. 175-197.

Stern, Ch., 2010 Clustering of environmental data using a Multi-agent System. In Proceedings of 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal, 2010.

Stern, Ch., 2009 Konzept zur Integration von Umweltdaten in kleinmaßstäbige DLM mit Hilfe eines Multi-Agenten-Systems. In Bundesamt für Kartographie und Geodäsie (ed.), Mitteilungen des Bundesamtes für Kartographie und Geodäsie. Arbeitsgruppe Automation in Kartographie, Photogrammetrie und GIS. Tagung 2008, (vol. 43). Frankfurt am Main, 2009, pp. 113-124.

Werder, S., 2009 Formalization of Spatial Constraints. In Proceedings of 12th AGILE Conference on GIScience, Hannover, Germany, 2009.