Deriving constraints for the integration and generalization of detailed environmental spatial data in maps of small scales

Christian Stern[1], Monika Sester[2]

[1]Faculty of Information Management and Media, 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

1. INTRODUCTION AND RELATED WORK

Towards an automated generalization workflow in cartography, the necessity and use of constraints as means to control the whole process is widely recognized. There have been many works that formulate or propose constraints or put them to use. For example, Beard (1991) proposes the use of four types of constraints: graphic constraints, structural constraints, application constraints and procedural constraints. Harrie (1999) presents a method for automatically displacing vector data using constraints, and the AGENT project (Lamy 1998) and all the following research works that make use of the AGENT paradigm to generalize topographic maps are based on the use of constraints.

Our current research focus lies on the extraction of structural constraints for the purpose of integration and generalization of spatial environmental data of high resolution in maps of lower scales. Spatial environmental data is displayed on maps in different shapes. It may be shown as point features (e.g. radioactivity measurement points or natural monuments like an old tree), linear objects (e.g. hydrography network), planar items (e.g. nature protection areas) or continuous surfaces (like wind velocity). Our research concentrates on the planar features, on protection areas in general and especially in water protection areas.

The International Union for Conservation of Nature and Natural Resources (IUCN) defines protected areas as “a clearly defined geographical space, recognised, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values” (Dudley 2008). For cartographers determined to integrate the data in maps of lower scale the definition’s term “clearly defined geographical space” is the most important part. According to Dudley (2008) this “implies a spatially defined area with agreed and demarcated borders. These borders can sometimes be defined by physical features that move over time (e.g. river banks) or by management actions (e.g. agreed no-take zones)”. In the German state of Baden-Württemberg, where our research area is located, protected areas used to be delimited by objects recognizable on a topographic map of the scale 1:25000. Today, they are captured on basis of a cadastral map in the scale 1:5000 (LUBW 2006) and the old geometries have been replaced step by step.

In Germany, water protection areas can be constituted to protect water bodies from lasting exposures on behalf of present or future water supply, to enrich ground water or to prevent harmful drainage of precipitation water and washing off and entry of soil components, fertilizers or pesticides into waters (WHG 2009, §51). Usually, water protection areas are divided into three protection zones. The first water protection zone is the catchment of spring and protects the well and its close range (10 to 20 meters). The second zone is the so called “close protection area” to secure the drinking water from contamination with bacteria. The flow time from the outer border of this zone to the well should last 50 days minimum. The “extended protection area” that covers the whole catchment area and where bans and use restrictions (e.g. ban to deposit waste) are in place is the third zone. Zones two and three may also be separated into subzones. (Wikipedia 2013). According to UVM (2003), water protection areas are technically demarcated on basis of a hydrological report. But because the hydrological border is not visible on earth’s surface, it has to be transformed in a formal process to a lawful border whose course is exactly defined and can be perceived at the location. Therefore, cadastral borders that are distinguishable in nature, or roads, paths, rivers, forest rims and other topographical objects or lines and markings that can be identified in nature and on a map at the same time are suitable for this task. It is also valid to use distances to aforementioned objects to describe the border of a water protection area (WPA), like “100 meters south of path xy”. On rare occasions, the border may be demarcated using signs or boundary markings in nature. An example of a water protection area (WPA) is shown in Figure 1a. Protection zones two (yellow crosshatches) and three (blue hatches) of WPA “Kaiserwald” can be distinguished. The outline segments are colored green and red (explanation follows further down). A cadastral map is used as background. There we can find out that the WPA is bordered by a street to the Northeast, a motorway and a moat (see also detailed map in Figure 1b) to the East, and forest, agricultural land and grass land in the South and West.

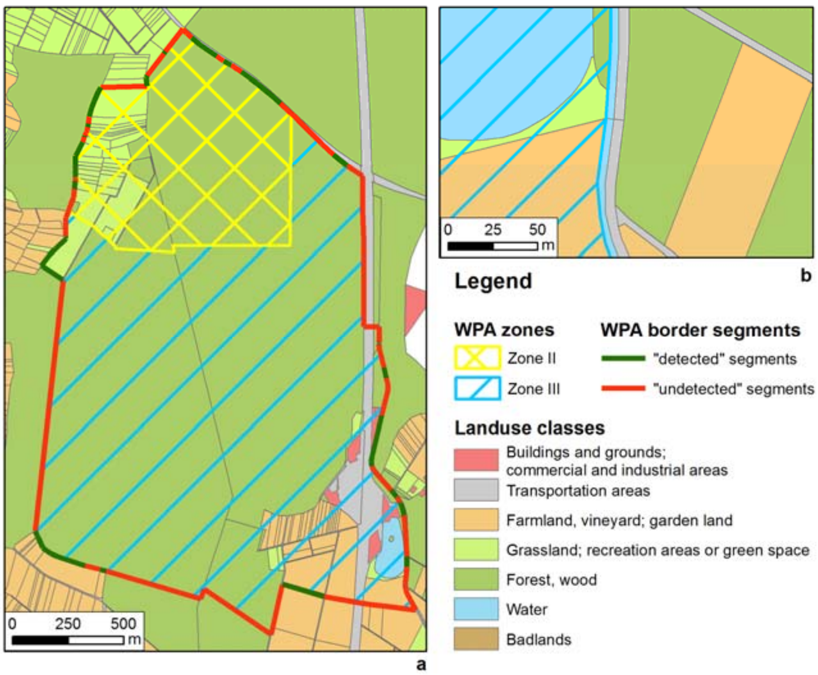

Figure 1. WPA Kaiserwald shown on top of a cadastral map (landuse classified as shown in Table 1). (a): The close protection area (yellow crosshatches) and the extended protection area (blue hatches) of the WPA can be distinguished. The colors of the WPA border segments show the result using the approach of Stern & Sester (2011). Dark green segments have been detected by their algorithm, while red ones have been omitted. (b): The detailed map reveals that the WPA’s border runs along a moat in its eastern part. Data source: LUBW, LGL.

Figure 2. WPA Kaiserwald displayed at small scales. Left: Print version, original scale 1 : 350 000. Right: Online version, original scale 1 : 250 000 (here both enlarged to approx. 1 : 140 000). Maps from LUBW, LGL. (Source: LUBW 2013a).

Table 1. Landuse classes after LUBW (2013b).

#	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		

In their paper, Stern & Sester (2011) outlined the need for displaying planar environmental spatial data on maps of lower scales and the graphical problems that arise by using this data that has been captured in high resolution. Figure 2 illustrates this problem showing WPA Kaiserwald displayed as a print version (left) and an online version (right) in small scales. In comparison to Figure 1a and 1b it can be seen that the original spatial relationships have not been preserved. The touch-relation of the motorway and the moat has changed to an intersect-relation for the street and a “does not touch-relation” for the latter. To overcome such shortcomings the above mentioned authors presented an approach to extract (structural) constraints that might serve as rules for integration and generalization of this data for use with digital landscape models of lower scale and for displaying it on small scale maps. They based their research on the fact that a water protection area is assumed to be outlined on the basis of parcel borders. By looking at the parcels and their landuse to the right and left side of the water protection area border they tried to find relationships between them and their surroundings that might be forged into rules which can be applied in the processes of generalization and integration. The dark green line segments in figure 1a show the parts of the WPA that were included in their analysis. The red ones could not be “detected” by their algorithm, meaning that these geometries do not have a touch relation with underlying parcel boundaries. The outcome of their approach was a list of frequencies how often a WPA is bordered by a certain combination of landuse classes, as classified by LUBW (2013b) which can be seen in Table 1.

Stern (2012) extended their approach on a wider investigation area and also included nature conservation and landscape protection areas in his research. He showed that for different protection area types like nature protection areas and water protection areas no common structural constraints can be found, only similarities if at all.

2. EXTRACTING CONSTRAINTS INCLUDING DISTANT GEOMETRIES

In our present work, we extended the approaches described above. The previous works only took in consideration the protection area border segments that had a "touch-relation" with directly underlying parcel boundaries. However, in many cases the data captured at the scale 1:5000 is not as accurately digitized as it is supposed to be or has a "built-in" distance because of demarcation issues (explained above). Often we found an offset between the WPA's and the cadastral data boundaries, ranging from a few centimeters to several meters at times. Hence we needed to extend the approach to "catch" cadastral parcel border segments that are "similar" and nearby to the WPA border segments but not incidental. Then we can assume that these found line segments are the "correct" outlines and use them for our further analysis.

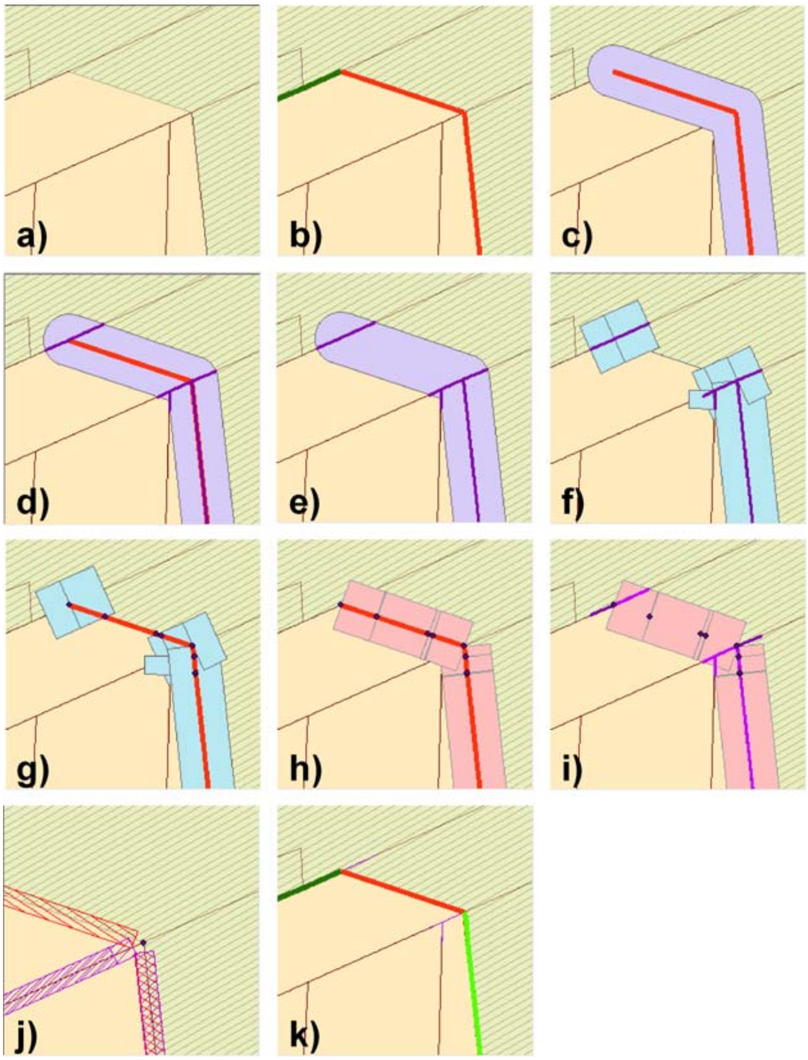

Figure 3. Processing steps to match water protection area boundaries to nearby parcel borders.

There are many works that tackle the problem of matching geodata sets and data integration. For example Walter & Fritsch (1999) presented a statistical approach to match two geographical data sets. More recently, Dahinden et al. (2013) and Dalyot et al. (2012) reported on geometrical methods for adjustments of vector data sets. For our use, we looked for an approach easy and fast to implement as we only needed the semantic information of the landuse class combinations rather than the exact geometric matches. Figure 3 illustrates the process steps that we used to find parcel border segments that corresponded to the WPA outline segments. In a) the initial situation can be seen. The excerpt of a map (screenshot taken at a scale of roughly 1:300) shows a part of a water protection area (WPA) in a light green color with hachures. In the background we can distinguish a cadastral map in a beige color, the parcel boundaries in brown. As we notice, the boundary of the WPA follows the parcel borders except for one parcel in the middle of the map piece that is crossed over. At b) the results of the previous approaches (Stern & Sester 2011) are outlined. In green we see a WPA boundary segment that could be matched with parcel boundary segments. The crossing segment in the middle obviously could not be matched to any border, but surprisingly the segment that seems to follow the parcel outline could not be matched either. Here a small shift occurs (in this case about 4 millimeters) that prevents a correct matching. To correct the erroneous data we set up the process that is shown in c) to k). At first, we created a search buffer (light purple color) around the WPA segments that previously had not been paired (c). Then we extracted potential matching candidates (dark purple colored line) by clipping the parcel outline segments with the search buffer (d) and (e). In the next step we buffered the potential candidates with the search distance (f) and calculated the intersection points (purple colored dots in g)) of these buffer polygons with the WPA segments. Now we could split the latter at these intersection points (due to technical limitations we had to create little buffers of one millimeter around the points that could be erased from the WPA segments rather than splitting them). This ensures that we have WPA segments and parcel outline segments that nearly have the same length if they run in parallel. In an intermediate step not to be shown in an image we introduced Universally Unique Identifiers (UUID) for each WPA segment and each potential matching candidate to be able to distinguish them later. Moving to step h) we buffered the WPA segments with the search distance. This enabled us to retrieve the final matching candidates by intersecting them with the buffers created previously, thus having now the two different UUIDs (one from itself and one from the corresponding WPA segment) for each of them (i). To be able to calculate shape measurements we once again created buffers around all segments with a minimum buffer distance of one millimeter (j). Now we could compute several measurements for each of the segment buffers, namely direction of the main axis, area, perimeter, perimeter^2/area ratio for both types and the distance to the corresponding WPA segment for the candidates. From these measurement values we calculated a score between zero and one that was somewhat reduced by a “distance weight”, taking away increasingly more if the distance between the WPA segment and candidate “grew”. A final score of near “zero” indicated “no match”, while a score near to “one” made us assume that we had found a match. So finally we looped through the complete dataset and looked for each WPA segment for the best match among the candidates. Then, in the last step (k), we replaced the WPA segment geometry by the one of the best match (light green colored line).

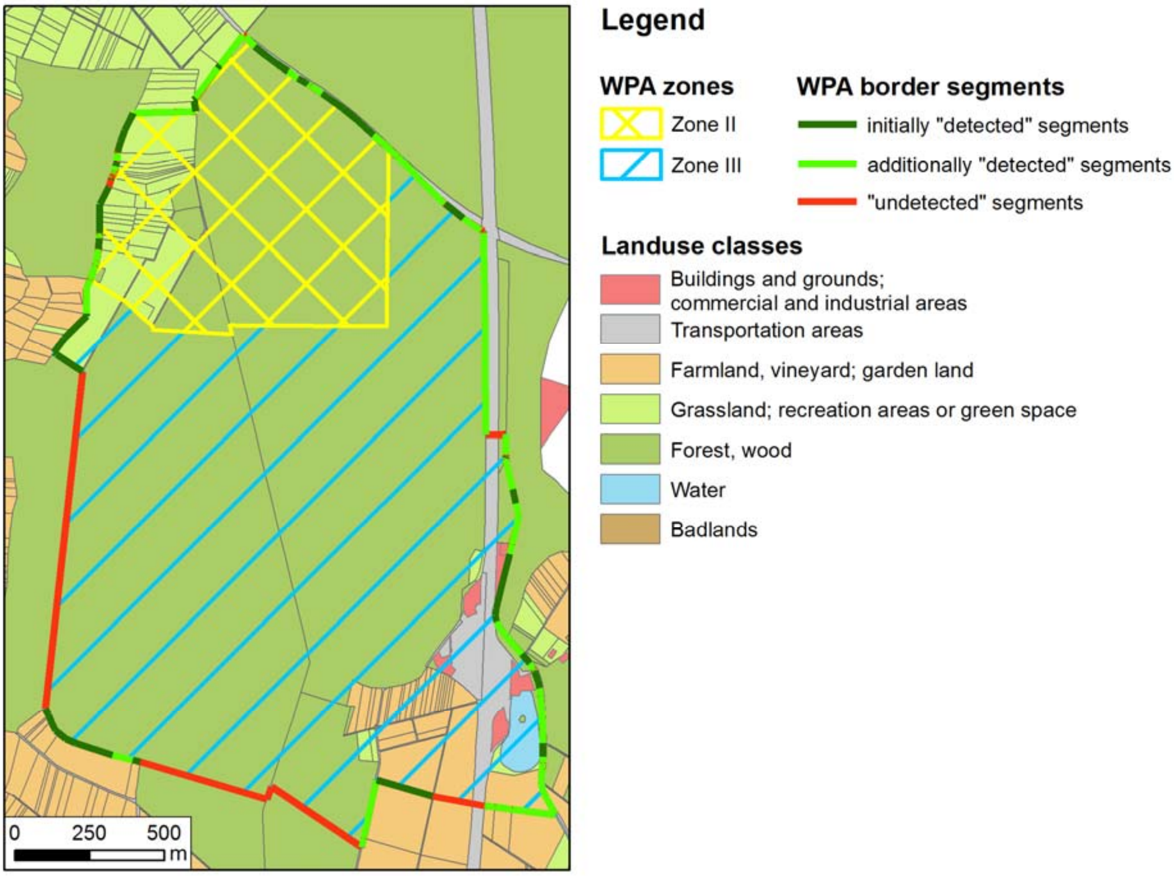

Figure 4. The border segments of WPA Kaiserwald after running the matching process with a search distance of 2 meters. The light green segments originating of the cadastral dataset could be additionally included in the data basis for the subsequent analysis. Compare to Figure 1 to see the changes. Data source: LUBW, LGL.

3. RESULTS

Using these processing steps, we could increase the matches found in our study area (district Emmendingen, Baden-Württemberg, Germany; as in Stern & Sester (2011)) from about 56 percent without a search distance to nearly 65 percent using a search distance of two meters, for example. Figure 4 illustrates the matching results (light green color) of WPA Kaiserwald as an example for the complete process. We figured out that the biggest increases are to be found among the low search distances, diminishing as the distance grows. For us, it seems that the optimal search distance is around two meters, comparing the results and the processing time.

Although improving the data basis for the subsequent analysis, in comparing the results of Stern & Sester (2011) using no search distance with our current results (search distance of two meters), we do not see much difference in the data. Table 2 opposes the 10 most important landuse combinations and their share of the WPA boundary length that have been found in the initial work to our current results.

Table 2. Comparison of the ten most frequent landuse combinations found in previous work by Stern & Sester (2011) (without search distance) with the results of the current work (2 meters search distance).

Landuse Class 1	Landuse Class 2	Share of total length in % (0 m search dist.)	Share of total length in % (2 m search dist.)
farmland, vineyard	transportation areas	20,70	21,66
forest, wood	forest, wood	20,04	18,49
farmland, vineyard	farmland, vineyard	8,41	8,37
transportation areas	forest, wood	5,14	5,04
grassland	transportation areas	4,25	4,49
grassland	grassland	3,93	3,90
badlands	transportation areas	3,89	3,74
buildings and grounds	transportation areas	3,80	3,96
grassland	forest, wood	3,49	3,51
transportation areas	transportation areas	3,38	3,64

Table 3. Comparison of the landuse classes found inside or outside next to boundary segments of water protection areas in the study area.

Landuse Class	Share of total length in % (INSIDE) (2 m search dist.)	Share of total length in % (OUTSIDE) (2 m search dist.)
forest, wood	26,8	23,09
transportation areas	26,13	23,52
farmland, vineyard	21,56	25,78
grassland	9,67	11,45
badlands	4,86	5,21
water	4,31	2,03
buildings and grounds	3,03	4,72
garden land	1,2	1,5
railway areas	0,96	0,74
recreation areas or green space	0,91	1,09
commercial and industrial areas	0,55	0,84
heathland, moor, swamp, fallow land	0,02	0,03

This time we did not only look at the frequency of landuse combinations in general, but tried to find out which of the landuse class types can be found inside or outside at the border of a WPA (see Table 3). Quite surprisingly we have an almost even distribution of the landuse classes between inside and outside of the protection area. One might have expected to find a “strong” constraint like that a certain landuse class (e.g. forest, wood) is more often to be found inside a WPA, while another class like transportation areas is likelier to be found outside. Nevertheless we can create a ranking table that can be used to control the generalization process. In Table 4 one can see the resulting ranking table from our test data. It shows the landuse class combinations that can be found at the borders of WPAs in our test area and their percentage of the total border length, focusing also on “inside” and “outside” the protection area. Ordering the percentage values with the highest number in the first row of the table, we could introduce ranks (as shown in the last column). Combinations with identical landuse classes were given no rank, as they deliver no useful information for controlling the generalization process. Figure 5 illustrates how such a table can help making decisions in the process. On the left side we see the initial situation. It is not clear if the WPA’s border should be adapted to the common border of the forest (F) and the street (s), the street and the grassland (G) or the river (r) and the grassland. Using the ranks of Table 4, we can solve the situation (right part of Figure 5) and generalize the WPA’s geometry to match the border of the forest (inside) and the street (outside). This tuple has the highest rank (position 3) of all available combinations.

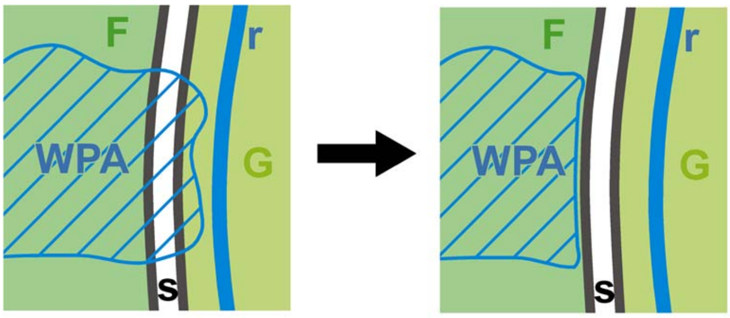

Figure 5. Example for a generalization process of a WPA using a ranking table. Left: Initial situation before generalization. Right: Situation after the generalization process that has been controlled by the results of a ranking table.

Table 4. Ranking table derived from the percentage of the total length of landuse class combinations found inside and outside border segments of WPAs in the study area.

Landuse Class 1 (INSIDE)	Landuse Class 2 (OUTSIDE)	Share of total length in % (2 m search dist.)	Rank
forest, wood	forest, wood	18,49	-
transportation areas	farmland, vineyard	11,99	1
farmland, vineyard	transportation areas	9,67	2
farmland, vineyard	farmland, vineyard	8,37	-
grassland	grassland	3,90	-
transportation areas	transportation areas	3,64	-
forest, wood	transportation areas	3,06	3
transportation areas	buildings and grounds	2,61	4
transportation areas	grassland	2,42	5
forest, wood	grassland	2,12	6

4. CONCLUSION

Our current work tried to improve the results that have been derived previously. We presented an approach to extract constraints for generalizing water protection areas by matching their boundary segments to underlying cadastral boundary segments even if they are not completely incident. By looking at the landuse classes that appear “inside” and “outside” the water protection area borders, we could see that they are almost evenly distributed. This will make it difficult to formulate “strong”, “general” constraints based on ranks that can be used for generalization or integration purposes. Nevertheless, if we derive “local” constraints that may only be applied for a specific WPA, we are sure that this will help well controlling the generalization process.

ACKNOWLEDGMENT

Finally, we want to thank the State Agency for Environment, Measurements and Nature Protection of Baden-Württemberg (LUBW) for supplying the data needed for this research work.

REFERENCES

Beard K, 1991, Constraints on rule formation. In: Buttenfield, B. and McMaster, R., (eds.), Map Generalization: Making Rules for Knowledge Representation, pages 121–135. Longman, London.

Dahinden T, Boljen J, Ruhe N, Dalyot S, Schulze J, Sester M, 2013, Geodatenintegration – Semantischer und geometrischer Abgleich von Vektordaten. In: zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement 138 (1), pp. 85-94, 2013.

Dalyot S, Dahinden T, Schulze M J, Boljen J, Sester M, 2012, Geometrical adjustment toward the alignment of vector databases. In: ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences I-4, pp. 13-18.

Dudley N (ed.), 2008, Guidelines for Applying Protected Area Management Categories. Gland, Switzerland: IUCN. x + 86pp.

Harrie L, 1999, The Constraint Method for Solving Spatial Conflicts in Cartographic Generalization. In: Cartography and Geographic Information Science, Vol. 26, No. 1, pp. 55-69.

Lamy S, 1998, The AGENT Project: Automated Generalisation New Technology: Proceedings of the 4th EC-GIS Workshop (Published by JRC).

LUBW (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg) (ed.), 2013a, Daten und Kartendienst der LUBW. URL: http://brsweb.lubw.baden-wuerttemberg.de/brs-web/index.xhtml. (Accessed: 07-22-2013).

LUBW (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg) (ed.), 2013b, Metadaten. URL: http://ripsmdk.lubw.baden-wuerttemberg.de/technik/alk/nutzung.htm. (Accessed: 04-07-2013).

LUBW (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg) (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).

Stern Ch, 2012, Bestimmung von Zwangsbedingungen für die Integration von Umweltdaten in kleinmaßstäbige DLM. In: Kartographische Nachrichten, 6/2012.
Stern Ch & M Sester, 2011, Towards extraction of constraints for integrating environmental spatial data in digital landscape models of lower resolution – a work in progress. In: Proceedings of 14th ICA/ISPRS Workshop on Generalisation and Multiple Representation, 2011, Paris.
UVM (Ministerium für Umwelt und Verkehr Baden-Württemberg) (ed.), 2003, WAABIS- Informationssystem Wasser, Abfall, Altlasten, Boden im Umweltinformationssystem Baden-Württemberg (UIS BW). WAABIS-Regelwerk zur Geodatenführung. Anleitung 3: Digitalisierung von Wasserschutz-, Quellenschutz- und Überschwemmungsgebieten. (Unpublished).
Walter V & D Fritsch, 1999, Matching Spatial Data Sets: a Statistical Approach. International Journal of Geographical Information Systems (IJGS), 13(5), pp. 445-473.
Wikipedia, 2013, „Wasserschutzgebiet“. In: Wikipedia, Die freie Enzyklopädie. 04-03-2013, 11:53 UTC. URL: http://de.wikipedia.org/w/index.php?title=Wasserschutzgebiet&oldid=116782232. (Accessed: 04-06-2013)
WHG, 2009, Gesetz zur Ordnung des Wasserhaushalts (Wasserhaushaltsgesetz - WHG). In BGBl. I p. 2585.