CLASS MODELLING OF BIOTOPE COMPLEXES – SUCCESS AND REMAINING CHALLENGES

D. Tiede [a]*, S. Lang [a b], F. Albrecht [a] and D. Hölbling [a]

[a] Centre for Geoinformatics, Salzburg University, Hellbrunnerstrasse 34, 5020 Salzburg, Austria
(dirk.tiede, daniel.hoelbling, florian.albrecht)@sbg.ac.at; stefan.lang@tu-berlin.de

[b] Institute of Landscape Architecture and Environmental Planning, Technical University of Berlin, Strasse des 17. Juni 145, 10623 Berlin, Germany

KEY WORDS: Object-based image analysis, aggregated functional units, CNL, hybrid approach

ABSTRACT:

Complex real world phenomena require means for reducing and modelling the underlying complexity. They ask for methods and tools to exploit the crucial information intelligently, for addressing dynamics and monitoring constant changes. The project Biotope Information and Management System (BIMS), carried out in a 3654 km² sized area in the region of Stuttgart/Germany, provides means for delineating and monitoring aggregated spatial units relevant for regional planning tasks. The project area is characterized by high dynamic growth and steady economic development. In the first phase of the project, we developed an adaptive per-parcel approach for delineating elementary units using SPOT-5 MS data (5 m GSD). In this paper we discuss our experiences from the second phase of the project, in which we developed algorithms for semi-automated class modelling of so-called biotope complexes based on these elementary units. The biotope complexes should represent highly aggregated, partly spectrally heterogeneous, but functionally homogenous units, following a mapping key provided by the Baden-Wuerttemberg Institute for Environmental Protection covering 19 different types of biotope complexes.

Due to their very heterogeneous structure, some classes could not be successfully modelled and therefore not derived in a fully automated way. For that reason we developed a tool (biotope complex modeller - extension for ArcGIS) to support experts in aggregating the remaining units in a comfortable and fast manner. New research developments in the field of object-based accuracy assessment which emerged from this project will be discussed.

1. INTRODUCTION

Many European metropolitan areas are characterized by a high dynamic and steady economic development and thus by spatial expansion and transformation of land use. As a consequence natural resources and particularly landscapes and their ecosystems face severe pressure. With a population density of 729 inhabitants per km² the Stuttgart region shows a significantly higher value than other densely populated areas like Munich or Frankfurt. The *Verband Region Stuttgart*, an association of local authorities within the Stuttgart region, elaborates an update of a regional landscape structure plan, which aims at determining the regional interests (status-quo, goals, and measures) of nature conservation and its management. One of its most important functions is to incorporate objectives of spatial planning. Decisions on this planning level require high quality and spatial exhaustive datasets making the *Verband* looking for an innovative, yet operational solution. As an answer, a project called Biotope Information and Management System (BIMS) provided a new, full-coverage geometry, with a scale-conform classification scheme (Schumacher & Trautner, 2006). The project agreed on using so-called biotope complexes, which represent highly aggregated, partly spectrally heterogeneous, but functionally homogenous units with 19 different classes. The methodological aim of this project was to develop a methodologically sound, yet operational and transferable approach to semi-automatically delineate and classify these biotope complexes using object-based image analysis (OBIA) for adaptive parcel-based segmentation class modeling (Lang et al., 2007).
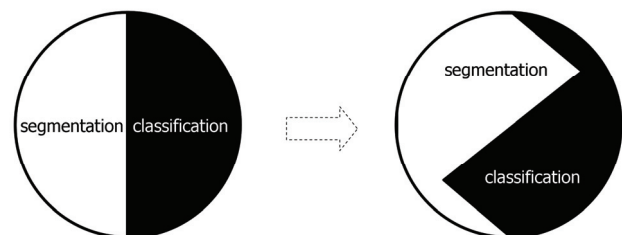

Fig. 1. Class modeling: from distinct realms of segmentation and classification towards an interlinked concept (from Lang, 2008, modified)

From a methodological point of view OBIA rests upon two interrelated methodological pillars, i.e. (1) segmentation / regionalization for nested, scaled representations; (2) rule-based classifiers for making explicit the required spectral and geometrical properties as well as spatial relationships for advanced class modeling. The term 'image analysis' indicates

---

* Corresponding author.

that the process of OBIA is iterative (cyclic) rather than linear and strict subsequent (cf. figure 1). By this, the process is highly adaptive and open for accommodating different categories of target classes, from specific domains, with different semantics, etc. (Lang, 2008)

2. STUDY AREA, DATA AND PRECEDING WORK

The study site is located in the south-western part of Germany, in the federal state Baden-Wuerttemberg, covering the region of Stuttgart with an area of 3654 km² (cf. Figure 2). Typical for this region is the heritage system of land ownership which has led to several iterations of splitting and dividing over the centuries. This practice has generated small parcels of real estate which are reflected in the Automated Cadastral Map (ALK). This digital cadastre from 2005 and partly from 2004 was used as the predefined target geometry. The actual information on land use and land cover was derived from a mosaic of four multispectral SPOT-5 scenes in pan-sharpen mode with 5 m ground sample distance (GSD), recorded between September 2nd and 6th, 2004. As a pre-processing step, orthorectification was done by use of a 5 m digital elevation model (DEM) with an orbital pushbroom model, implemented in the Leica Photogrammetry Suite. For co-registering an existing mosaic of orthophotos with 0.25 m GSD was used. Due to the use of very accurate ground control points (GCPs) during co-registration process the resulting data shows a high spatial accuracy with a low root mean square error of displacement (considerable below 1 pixel). At last the SPOT-5 data was clipped to the boundaries of the administrative districts regarding a 500 m buffer.
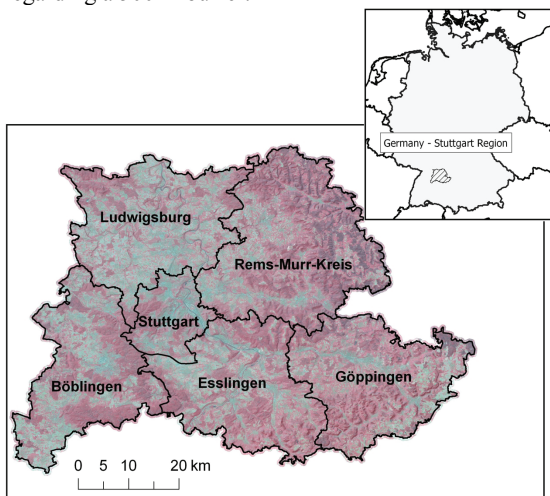

Fig. 2:   SPOT-5 mosaic covering the Stuttgart region (from Tiede et al. 2007)

In the first phase of the project we focused on the delineation of individual objects and developed an adaptive per-parcel approach (cf. Lang et al. 2007, Tiede et al. 2007). By applying our approach it was possible to overcome the entailed problems of the fine-scaled cadastre data like redundant boundaries within homogenous land cover types or inadequate boundaries which do not reflect the ones needed. This was done by merging, splitting and cleaning of objects (Lang et al. 2007).

Starting from these elementary units, representing (spectrally) homogenous land cover types, the final aim, as indicated above, was to derive biotope complexes. These biotope complexes represent highly aggregated, partly spectrally heterogeneous,

but functionally homogenous units, following a mapping key provided by the Baden-Wuerttemberg Institute for Environmental Protection covering 19 different types of biotope complexes (LfU, 2003, see Table 1). The mapping key lists several quantitative interpretation cues like minimum and maximum size of potential biotope complexes, proportions of different types of land use, e.g. grassland / cropland, percentage of open orchard meadows or tree species composition in the forest types. Additionally, some rather vague and qualitative specifications are given, e.g. a compact shape for biotope complexes.

	Biotope complex	Minimum mapping area
I	Residential buildings of cities, towns and villages	0.5 ha
II	Mixed use development, industrial or commercial areas	0.5 ha
III	Vehicular and pedestrian networks minimum width	20 m
IV	Green space, cemeteries, recreation areas and domestic gardens	0.5 ha
V	Mineral extraction sites	2 ha
VI	Infrastructure and storage area	0.5 ha
VII	Arable land, poor in accompanying habitat structures	4 ha
VIII	Arable land, rich in accompanying habitat structures	2 ha
IX	Vineyards and fruit plantations	2 ha
X	Special cultivation area	2 ha
XI	Mixed arable land and grassland area	2 ha
XII	Agriculturally improved grassland	2 ha
XIII	Extensive grassland	2 ha
XIV	Open orchard meadows	2 ha
XV	Abandoned open areas	2 ha
XVI	Fens and bogs	0.5 ha
XVII	Broad-leaved forests	5 ha resp. 2 ha
XVIII	Coniferous forests	5 ha resp. 2 ha
XIX	Areas of water	0.5 ha

Table 1:   Types of biotope complexes according to LfU (2003, from Schumacher and Trautner, 2006, modified)

Auxiliary information layers were integrated providing valuable input for the modelling process: (1) Urban areas and main roads were taken from the administrative topographical cartographical information system (ATKIS) and masked out in the modelling process, (2) biotope information from nature conservation mapping and (3) water bodies and open orchard meadows from ATKIS.

3. METHODS

3.1 Class-modelling

The approach was realised by developing rule-sets using Cognition Network Language (CNL) within the Definiens Developer Environment. CNL, similar to a modular

programming language, supports programming tasks such as branching, looping, and defining of variables. More specifically, it enables addressing single objects and supports manipulating and supervising the process of generating scaled objects in a region-specific manner. By this, the process steps of segmentation and classification can be coupled in a cyclic process, what we call class modelling (Tiede et al., 2008).

A basic land use / land cover multi-scale classification of the created elementary units draws upon the spectral properties, which can be selected for each unit in aggregated form. As gradual transitions occur in this process, in some cases fuzzy-classifications were applied, implementing labelling based on probabilities. Further, by means of spatial relations, the additional information layers could function as hints and were incorporated into the classification.

Based on this first classification we attempted to model biotope complexes, according to the LFU (2003) guidelines. Figure 3 gives a schematic overview of the modelling process. The iterative cycles of our class modelling approach are, in this case, a crucial element for stepwise growing and merge operations of objects. Otherwise specifications for the shape or the proportion of land use types for the biotope complexes could not be controlled in an adequate manner.
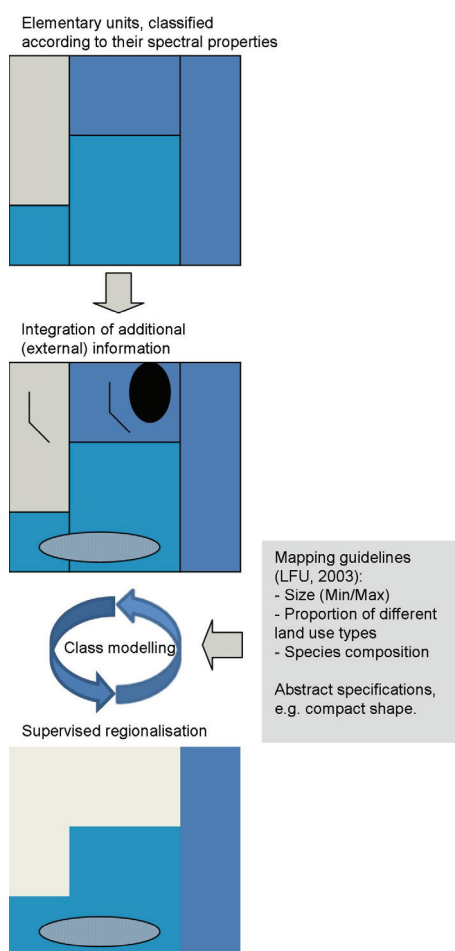

Fig. 3: Schematic overview of the biotope complex modelling process

3.2 Biotope complex modeller – Extension for ArcGIS

A hybrid approach was chosen in cases, in which the results of the semi-automated generation of biotope complexes were not satisfying due to a high number of degrees of freedom and the limitation of the mono-temporal satellite data set. This was particularly true for complexes showing spectral heterogeneity such as mixed arable land and grassland area (XI) but also for the complex types arable land (VII and VIII) and agriculturally improved grassland (XII). In these areas a processing time and effort optimising method was developed by coupling specific strengths of the evaluation systems "man" and "machine" (Lang et al., 2008). A tool was programmed (biotope complex modeller - Extension for ArcGIS, Figure 4) to support experts in aggregating the remaining units in a comfortable and fast manner. The interpreter instantly obtains information on selected units, aggregated area and land use composition. Thus the human perception capability was efficiently integrated. The implementation of this hybrid approach (Lang et al., 200) represents a major achievement in the area of semi-automated supply of complex target classes.
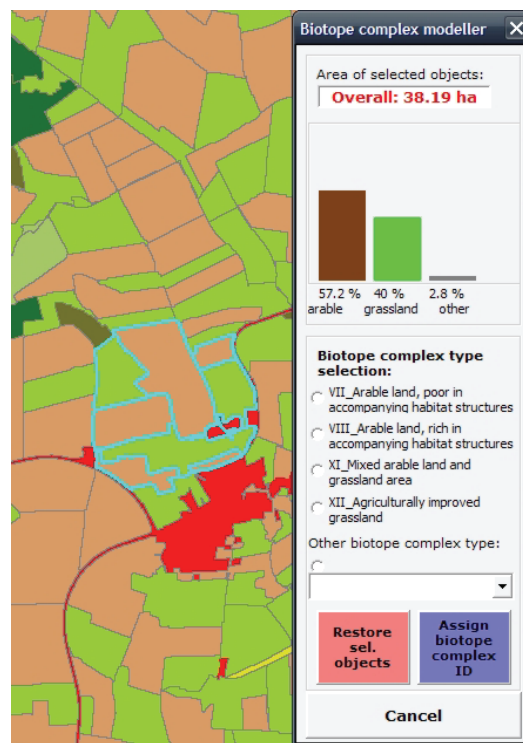

Fig. 4: Biotope complex modeller, extension for ArcGIS, supporting experts in aggregating classified elementary units in a comfortable and fast manner.

4. RESULTS AND ACCURACY ASSESSMENT

Altogether 31,698 biotope complexes were delineated for the whole Stuttgart Region (Figure 5 and 6). This figure includes complexes aggregated by experts using the above mentioned hybrid approach. The average size of each complex was 11.5 ha. An evaluation of the generated units was carried out by experts for habitat mapping, who did a qualitative mapping of each biotope complex (e.g. the potential importance for target species). If a biotope complex was not adequately modelled, it

has been manually changed accordingly. We conducted comparative analysis of the revised result with the modelled biotope complexes for a sub-area of 583 km². Therefore about 86 % of the biotope complex boundaries were correctly delineated and about 96 % of the removed boundaries (starting from cadastre data level) were correctly removed (cf. Tiede et al. 2007).
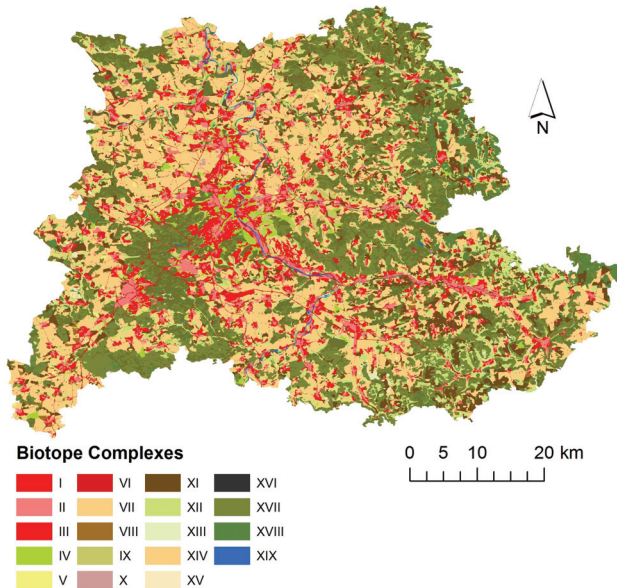

Biotope Complexes

Fig. 5: Resulting biotope complexes for the whole Stuttgart Region including complexes aggregated by experts using the hybrid approach.

Fig. 6: Successfully derived biotope complexes (highlighted, with underlying orthophotos): (1) automatically delineated open orchard meadows (above) (2) mixed arable land and grassland area using the proposed hybrid approach

The classified elementary units were evaluated using a statistical point-based accuracy assessment. Units with a

classification mainly based on additional information layers (urban areas, main roads) were not taken into account. About 300 randomly distributed points (at least 20 per class) were generated and visual comparison with the SPOT imagery was performed. The overall accuracy was about 85 % ($K^\wedge$ = 0.8356).

5. OUTLOOK: OBJECT-BASED ACCURACY ASSESSMENT – NEW RESEARCH DEVELOPMENTS

The heterogeneous character of the biotope complexes, resulting from the described workflow shows that a pixel-based site-specific accuracy assessment is limited. The class descriptions overlap for some classes/biotope complexes. For instance biotopes are integrated (mixed) from other classes. For a site-specific accuracy assessment with sample points taken some classes cannot be evaluated. Their character is only revealed (i.e. the correct label) if considered as a whole object (due to their heterogeneity within themselves, e.g. an orchard consisting of trees and meadow- a site-specific assessment would address a different level of hierarchy).

Accordingly, a method of object-based accuracy assessment, Object Fate Analysis (OFA) has been proposed, which addresses objects as a whole (Schöpfer et al., 2008). An object consists of an area which is determined by its delineation. The thematic label is associated with the object area. The delineation as such is an additional property (an abstraction of the object) that is not evaluated by traditional site-specific accuracy assessment. The boundary determines the degree to which a classified object as a whole corresponds to a reference object used for evaluation. This has implications on the comparability of the thematic content of the two object representations. OFA was initially used for analyzing how specific landscape units changed from one point in time to another (Lang et al., 2008; Schöpfer et. al., 2008), and then exemplarily applied to the data sets used in the BIMS project (Albrecht, accepted). OFA is a combination of two different object comparison procedures. An extended epsilon error band for polygons is applied (Zhang & Goodchild, 2002), dividing the topological relations between classification objects and their corresponding reference objects into two clusters, 'similar to disjoint' (C1) and 'similar to equal' (C2) (Egenhofer & Herring, 1991; Straub & Heipke, 2004). Applying the concept of topology takes into account the fact that a resulting object may cover the same area as more than one reference object does and vice versa. Therefore, there is an n-to-m-relationship between classification objects and reference objects. Some of the objects' boundaries in the classification may not have a counterpart in the reference. So applying solely the error band method will not lead to satisfying results. The additional use of the topological comparison is required. (Albrecht, accepted)

In OFA the relation between each classification object and the corresponding reference object can be categorized into the categories 'good', 'expanding', 'invading' and 'not interfering', where the first and the last category are distinguished into two subcategories (figure 7).

For each reference object all classification objects are compared that do not have the relation 'disjoint' to it. This excludes the relations of the category 'not interfering I' from further analysis. As the objects in this relation do not share any area they cannot be called corresponding objects. The first major decision criterion is applied with a point-in-polygon request on

the classification objects. The topological relationships are distinguished into whether they belong to a greater extent to the interior (C2) or to the exterior (C1) of a reference object, which means the upper categories or the lower categories of figure 7, respectively. The second decision criterion is achieved with the help of a buffer. It is calculated to the inside for distinguishing the C1 relations and to the outside for distinguishing the C2 relations. If a classification object overlaps the reference object boundary only by a small distance and does not cross the buffer the overlap is considered to be insignificant and the relation will be assigned to the categories 'not interfering II' for C1 relations and 'good II' or 'good I' for the C2 relations. If the buffer is crossed this will result in the assignment to the categories 'invading' (for C1) and 'expanding' (for C2; Albrecht, accepted).
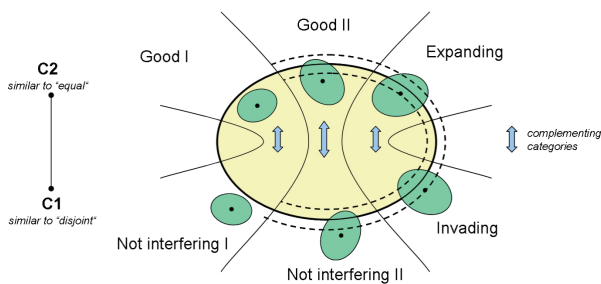

Figure 7: Categorization of the interaction between classification objects and reference objects (Schöpfer et al., 2008; modified)

This categorization process is applied several times with an increasing buffer distance. It is observed how the categories for all the relations change with buffer distance. A specific buffer distance can be found where many of the 'expanding' relations will have changed to 'good' relations. The ratio R_{e_ge} in equation 1 compares these relations for every buffer distance and can be used to evaluate this change.

$$R_{e\_ge} = \frac{n_{exp}}{n_{good} + n_{exp}}, \qquad (1)$$

where   n = the number of all relations of the specified category occurring for all comparisons between the classification objects and the reference objects (taken from Albrecht, accepted, after Lang et al., 2008).

These relations will occur for those classification objects, where the boundary is subject to random error only. Therefore the specific buffer distance can be associated with the error band (Albrecht, accepted).
The approach was tested using data from the BIMS project. We used preliminary classification results without enhancement by manual interpretation. As reference data the final BIMS product (after the qualified mapping of each biotope complex by experts for habitat mapping) was used. OFA was carried out and ratios were calculated for buffers from 0 to 10 m. At 3.5 m buffer distance a significant decrease of the ratio R_{e_ge} occurred, which then stayed quite stable for larger buffer distances. Half the diagonal of a SPOT pixel has a length of 3.54 m. This buffer distance therefore is related to the raster-to-vector-conversion within the data production process (ibid.). The results from the case study show that error present in the object delineation of an OBIA derived classification could be measured.

6. CONCLUSIONS

A class modelling approach has been successfully applied in the project Biotope Information and Management System (BIMS), providing means for delineating and monitoring aggregated spatial units relevant for regional planning tasks.

Class modeling provides flexibility in providing problem-oriented solutions for advanced analysis tasks, involving both operator and expert or user in rule set generation. As a production system it lacks any learning component (as opposed to neural networks or support vector machines), but instead it is capable addressing structurally aggregated classes and supporting scene-specific high-level segmentation with delineation of conceptual boundaries (Tiede et al., 2008; Tiede & Lang, accepted).

However, it has to be said that in terms of efficiency, multi-temporal data with higher spatial resolution would have lead to even more reliable results. In effect this would suggest a lower effort of post-processing as well as a higher degree of automation.

Using Object Fate Analysis it is possible to conduct a spatial accuracy assessment for object boundaries. There still needs to be more research on how this method can be enhanced for a real object-based accuracy assessment that also includes issues of thematic content.

REFERENCES

Albrecht F. accepted. Assessing the spatial accuracy of object-based image classifications. In: Car, Strobl, Griesebner (eds) Geospatial Crossroads, Wichmann Verlag.

Egenhofer MJ., Herring JR., 1991. Categorizing binary topological relations between regions, lines, and points in geographic databases. University of Maine, National Center for Geographic Information and Analysis, Orono, Maine, USA, 28 pages.

Lang S., 2008. Object-based image analysis for remote sensing applications: modeling reality - dealing with complexity. In: Blaschke T, Lang S, Hay GJ (eds) Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications. Springer, Berlin (pages pending)

Lang, S., Schöpfer E., Langanke T., 2008. Combined object-based classification and manual interpretation - Synergies for a quantitative assessment of parcels and biotopes. *Geocarto International*, 23 (4), pp 1-16.

Lang, S., Tiede, D., Schumacher, J. & Hölbling, D., 2007. Individual object delineation revising cadastral boundaries by means of VHSR data. In: M. Ehlers & U. Michel (Hrsg.), Remote Sensing for Environmental Monitoring, GIS Applications, and Geology VII conference. Vol.6749, 8 pages. Florence, Italy: SPIE.

LfU, Landesanstalt Für Umweltschutz Baden-Württemberg, Hrsg., 2003. Handbuch zur Erstellung von Pflege- und Entwicklungsplänen für die Natura 2000-Gebiete in Baden-Württemberg. Version 1.0. – Fachdienst Naturschutz, Naturschutz Praxis, Natura 2000: 467 pages; Karlsruhe.

Schöpfer E., Lang S., Albrecht F., 2008. Object-fate analysis – spatial relationships for the assessment of object transition and correspondence. In: Blaschke T, Lang S, Hay GJ (eds) Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications. Springer, Berlin (pages pending)

Schumacher J. & Trautner J., 2006. Spatial Modeling for the purpose of regional planning using species related expert-knowledge. The Biotope Information- and Management System of Stuttgart Region (BIMS) and its deduction from the Information System on Target Species in Baden-Württemberg, E. Buhmann, S. Jørgensen and J Strobl. (eds.), Trends in Knowledge-Based Landscape Modeling. Wichmann-Verlag, Heidelberg.

Straub BM., Heipke C., 2004. Concepts for internal and external evaluation of automatically delineated tree tops. IntArchPhRS. Band XXXVI Teil 8/W2, Freiburg, pp 62-65.

Tiede, D., S. Lang & C. Hoffmann 2008, Type-specific class modelling for one-level representation of single trees. In: Blaschke T, Lang S, Hay GJ (eds) Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications. Springer, Berlin (pages pending)

Tiede D. & S. Lang ,accepted. Distributed computing for accelerated dwelling extraction in refugee camps using VHSR satellite imagery. In: Car, Strobl, Griesebner (eds) Geospatial Crossroads, Wichmann Verlag, (pages pending)

Tiede, D., Moeller, M., Lang, S., & Hölbling, D., 2007. Adapting, Splitting and Merging Cadastral Boundaries according to homogenous LULC Types Derived from SPOT 5 Data; In: Stilla U et al (Eds) PIA07. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (3/W49A), pp 99-104

Zhang J, & Goodchild MF, 2002. Uncertainty in geographical information, London.