PRACE GEOGRAFICZNE, zeszyt 146

Instytut Geografii i Gospodarki Przestrzennej UJ
Kraków 2016, 7–18
doi:

ASSESSING ACCURACY OF FOREST COVER INFORMATION ON HISTORICAL MAPS

Urs Gimmi, Christian Ginzler, Matthias Müller, Achilleas Psomas

Abstract: Assessing the uncertainty of historical map information is a prerequisite for using this source type for land cover reconstructions. We assessed the accuracy of forest cover information from 1930/1940 Swiss topographic maps (Siegfried maps) by comparing them with forest cover gained from orthophotos and terrestrial photos from them same period. Orthophotos were considered to contain the true forest cover information and consequently inconsistencies between historical maps and orthophotos were considered as true errors (over- or underestimation of the forest cover on the map). Terrestrial photos reflect the surveyors perspective in the field and we therefore used this source type to develop hypotheses on potential reasons for inaccuracy of map information. Generalized linear models were used to test the hypotheses. The results confirmed our expectation that disagreement between maps is more likely to occur near forest edges, at higher altitudes and in less accessible and visible areas. Contrary to our expectations terrain slope did not play a significant role to explain inaccuracies surprisingly high rates of inaccuracy occurred in forest core areas.

Keywords: forest cover change, historical maps, aerial photographs, terrestrial photographs, map accuracy, Swiss Alps

Introduction

Long-term forest cover change datasets form an important source of information to address various ecological and societal questions. Major changes in land-use regimes often include a strong forest component. Probably the most prominent example is the forest transition theory which describes the shift from a net forest decrease to forest expansion observed in many regions and which is strongly linked to fundamental socioeconomic changes (Mather 1992; Rudel *et al.* 2005). Information on past forest cover changes has also been used to assess the effectiveness of protected areas (DeFries *et al.* 2005). Spatially explicit time series of forest cover help to separate old (ancient) from new forest areas and to study past dynamics of forest cover change (Munteanu *et al.* 2015). This information can be used as powerful indicators in ecological and biodiversity assessments (Graae *et al.* 2003; Herrault *et al.* 2016). Consequently, historical forest cover information is key in defining appropriate management and conservation strategies (Hermy *et al.* 1999).

However, to establish long-term time-series of forest cover changes is a challenging and time-consuming task. Satellite imagery is a well-established source to detect large-scale land cover changes yet they reach back only a few decades. Historical maps offer the opportunity to extend time-series to a few centuries. Many studies employed historical maps to reconstruct long-term land use/land cover changes (e.g., Gimmi *et al.* 2011; Skalos *et al.* 2011), and particularly to quantify changes in the extent of forest area (e.g., Kozak 2003; Skalos *et al.* 2012; Wulf *et al.* 2010; Kaim *et al.* 2016). In many cases, proper assessment of the accuracy of forest cover information extracted from historical maps is lacking (Leyk, Zimmermann 2004). However, an estimate of the uncertainty inherent in the map information is essential in order to adequately assess the accuracy of the forest cover change analyses, and to evaluate the opportunities and limitations of their interpretation.

Leyk *et al.* (2005) developed a typology of potential uncertainties inherent in historical map information. The authors distinguish between production-oriented (derive from data collection and map production), transformation-oriented (caused by data processing) and application-oriented types of uncertainty. In this study, we aim to assess the uncertainty of forest cover information in the alpine region of Switzerland derived from topographic maps from the 1930/1940s (so-called Siegfried maps)related to the most important production-oriented error type by comparing forest cover data generated based on historical map information with forest cover data extracted from independent sources (aerial and terrestrial photographs). Finally we discuss the effects of our findings on potential applications of past forest cover information.

In this study we address two main research questions:

– How accurate is forest cover information on Siegfried maps from the 1930/1940s?
– Are there systematic patterns in forest cover uncertainty that can be related to the topography, accessibility and to the shape of the forest/ non-forest landscape?

Data and methods

In this study we assess the accuracy of forest cover information on Siegfried maps available for entire Switzerland by comparing forest cover information from maps with independent sources of information from the same time period. As the first reference, we used information from 1940s orthophotos from five communities across the Swiss Alps (Giswil, Bürglen, Valendas, Lostallo and Lenk, Fig. 1). Forest cover information acquired from this source was considered as the 'true' forest cover information. An error map including information on different error types was then be generated by the spatial overlay of map and orthophoto information. As a second independent source we used forest cover information from terrestrial photographs taken at the same time. The fact that the Siegfried maps were produced based on terrestrial surveys made the terrestrial photographs particularly interesting for our study as they represented the surveyor's perspective in the field. Based on the qualitative comparison between orthophotos and terrestrial photos we developed a set of hypotheses to explain potential uncertainties in forest cover mapping. We then tested these hypotheses based on identified disagreement between forest cover on the historical maps and reference forest cover as displayed on contemporary orthophotos. The conceptual study design is illustrated in Figure 2.

Extracting forest cover information from maps

Siegfried maps were named after Colonel Hermann Siegfried, head of the Swiss Topographic Office in the second half of the 19th century (Gugerli, Speich 2002). These countrywide maps were published in several editions from 1870 to 1949 at scale 1: 25 000 for the lowlands and 1: 50 000 for the mountainous regions. We picked the last edition of this map series by compiling multiple maps into one composite picture for an area of roughly 10 000 km^2 covering five cantons in the eastern part of the Swiss Alps. The map dates range between 1917 and 1944, however more than 95% of the maps (including those used in this study) date from the period 1932–1942. The spatial accuracy of the digitized maps was evaluated by comparing intersection of coordinate lines with modern topographic maps and it was found to be 3 m in average and 20 m in maximum (Ginzler *et al.* 2011). The instructions for the surveyors how to indicate forests on the map were quite vague. Most importantly, the surveyors had to record all forest according to their density and forest edges had to be drawn either as clear boundaries or if there was no clear forest boundary without sharp outlines (this was often the case at the timber line). Further, avalanche tracks had to be drawn as special characteristics of mountain forests. For this study, forest cover information was extracted by manually digitizing forest patches from the Siegfried maps (example shown in Fig. 3a).

Extracting forest cover information from orthophotos

For the selected five communities (see Fig. 1) aerial imagery dated from the same time as the Siegfried maps was available. Images were purchased from the Federal Office of Topography (swisstopo). The images had an average scale of 1: 20 000. After orthorectification the orthoimages were used to extract forest cover information by vectorizing polygons for patches of group of trees and continuous forest patches and points for single trees in areas with less dense tree cover. In a post-processing step the individual trees and small forest patches were grouped together based on distance threshold and alpha-shaping radius (Boesch, Ginzler 2013; example shown in Fig. 3b).

Extracting forest cover information from terrestrial photographs

For the same five communities a set of 25 terrestrial photographs roughly dating from the same time as the local Siegfried maps have been collected from local and online photo archives. Requirements for suitable photographs were the following: (a) the photo needed to have good quality in terms of resolution and bit-depth; (b) the photo had to show a considerable part of the forested landscape; and (c) the forests visible had to be not too far in the distance and be viewed from an appropriate, not too flat angle (ideal would be perpendicular to view-direction). From the 25 collected historical photographs 10 have met these requirements (minimum 1, maximum 3 per community). Forest cover information from the selected terrestrial photographs was extracted using the Monoplotting tool GIS Suite (Bozzini *et al.* 2012). For each photograph the first step was to define the approximate camera location and at least five well distributed control points for georeferencing. Finally, forest cover was extracted by vectorizing forest patches as polygons on the terrestrial photographs. The resulting features could be exported to ArcGIS 10.2 for further processing (Fig. 3c and 3d).

Comparing forest cover information from different sources

Comparing Siegfried maps with orthophotos

In the first step, forest cover information based on historical maps was compared to forest cover based on the orthophoto interpretation (considered as true forest cover). We applied a 10 m buffer to the inside and to the outside for both the forest cover based on historical map and orthophotos in order to implement an accepted tolerance in disagreement of 20 m (considering the maximum positional error of the maps). The results of the spatial overlay of the forest cover information from both sources provided a map showing areas with agreement and disagreement between forest cover from the historical map and the orthophotos. From this map we were able to distinguish four different cases:

- true-negative (historical map = non-forest; orthophoto = non-forest): correct interpretation;
- true-positive (historical map = forest; orthophoto = forest): correct interpretation;
- false-negative (historical map = non forest; orthophoto = forest): underestimation by historical map (error type 1);
- true-positive (historical map = forest; orthophoto = non forest): overestimation by historical map (error type 2).

Comparing orthophotos with terrestrial photos

We defined areas which were well visible on the terrestrial photos (not too far in the distance and viewed from an appropriate, not too flat angle) because a comparison with the orthophotos was meaningful only for those areas. Instead of a full quantitative comparison we qualitatively distinguished between areas with particular good and bad agreement in order to formulate hypothesis on potential reasons for inaccuracy of map information. Terrestrial photos are an excellent basis to develop meaningful hypothesis because they reflect the surveyor's perspective in the field.

Formulating and testing hypothesis

We formulated a number of general hypotheses partly based on our observations when comparing terrestrial and orthophotos for variables related to topography, accessibility, visibility and the shape of the forest/non-forest landscape (Tab. 1) in order to test the effect of these variables on map accuracy.

Hypothesis based on topography

Our basic assumption was that forest cover on historical maps was more accurately depicted in lower altitudes and in flat terrain (both error types). As parameters we calculated the altitudinal difference to the lowest point in the community and the terrain slope. Both parameters were derived from the Swiss digital elevation model.

Hypothesis based on accessibility

As the historical maps were based on terrestrial surveys we assumed that map accuracy would be higher in well accessible regions. As an indicator for accessibility we used the distance: (a) to the road network; (b) to single houses not connected to roads; and (c) to fixed survey points used for the triangulation. All information

Table 1. Overview of explanatory variables for modelling forest cover accuracy

Type of variable	Variable	Source and processing
Topography	Altitude (difference to the lowest point of community)	DEM 25
	Slope	Derived from DEM 25
Accessibility	Distance to roads/houses/survey points	Vectorization from Siegfried maps and distance calculation
Visibility	Number of survey points a location is visible from (≤3000 m)	Siegfried maps and DEM 25
Shape of forest/ non-forest landscape	Edge/core/gaps of the forest and non--forest area	Forest cover information from orthophotos processed with morphological spatial pattern analysis (MSPA, Vogt *et al.* 2007)

was extracted directly from the historical maps.

Hypothesis based on visibility

We assumed that accuracy would be higher in regions that were well visible for the map surveyors from different locations. As an indicator for visibility we used the number of triangulation points a specific location is visible from (not further than in 3 km distance).

Hypothesis based on the shape of forest/non-forest landscape

We assumed higher agreement between historical maps and orthophotos in interior forest locations (core) and lower agreement close to forest edge (both error types). Particularly error prone are smaller forest gaps within forest patches because they are not well visible for the map surveyors from the ground (error type 2). As an indicator we analysed morphological spatial patterns applying the methodology developed by Vogt *et al.* (2007, Fig. 4), using three simplified morphology classes: core; edge (aggregation of edge, perforation, bridge, loop, branch and islet) and gap (background surrounded by perforation).

Model setup

Probabilistic, spatially explicit models were calibrated for the five communities. Generalized linear models (GLM; McCullagh, Nelder 1989) were used as the statistical technique. Sampling locations for model calibration and validation for the different error types were generated using a random stratified approach where all possible four cases described above (true-positive, true-negative, error type 1, error type 2) were included (Tab. 2).

Additionally sampling locations closer to 100 m from each other were removed.

To evaluate the predictive power of the model, we applied a split sample approach: the model was calibrated with 70% randomly chosen data and evaluated on the remaining 30%. This process was repeated 10 times and the model validation results were assessed using the true skill statistic (TSS) metric. TSS is a measure of agreement between predicted and observed values (Allouche *et al.* 2006). The measure is computed as specificity (fraction of correctly predicted presences) + sensitivity (fraction of correctly predicted absences) – 1, and varies between negative values (systematically wrong), 0 (random model) and 1 (perfect agreement).

Table 2. Number of sampling point per community and error type used in the statistical modelling

Community	Sampling points	
	type 1 error underestimation	type 2 error overestimation
Bürglen	446	1473
Giswil	648	284
Lenk	26	842
Lostallo	258	106
Valendas	505	295

Results

Overall agreement between maps and orthophotos

In all communities error type 1 (underestimation of forest cover) is much better predicted by GLM model than error type 2 (TSS 0.81 *versus* TSS 0.44 while the variability of model performances among communities is similar for both error types (difference of roughly 0.3 from highest to lowest TSS) (Tab. 3). Across all communities the GLM model shows a good agreement.

Factors explaining map inaccuracy

The shape of forest area (morphology) is clearly the dominating factor explaining error type 1 (Fig. 5). Regarding to the response curves (Fig. 6) most misinterpretations in the map occurred within the core forest area and at the forest edge (labelled as 1 and 3 in the response graph). Apart from morphology, only the distance to roads and houses contributes to the model. From the response curve we could assume that more type 1 errors occurred further than 1.5 km away from roads and single houses. All other variables do not contribute to the model.

Table 3. Overall agreement between maps and orthophotos for the two error types. TSS is used as a measure of the GLM accuracy.

Community	True Skill Statistics (TSS)	
	Type 1 error	Type 2 error
Bürglen	0.88	0.50
Giswil	0.82	0.57
Lenk	0.77	0.60
Lostallo	0.54	0.34
Valendas	0.74	0.30
overall	0.81	0.44

For the error type 2 the picture is more balanced with altitude being the most important variable and morphology (above all of non-forest area), distance to roads and houses as well as visibility playing a role (Fig. 5). Overestimation of forest cover on maps occurred most likely in mid to high altitudes, close to the forest edge (first 10 m outside of forest, that is edge of non-forest morphology), if the nearest road or single house was approximately one kilometre away and at locations that were visible only from a few survey points (Fig. 7).

Discussion and conclusions

Historical maps offer a unique opportunity to reconstruct past land cover in particular forest cover dynamics. However, as maps are an interpretation and a generalized representation of the real world there is always some uncertainty included in historical map information.

Our case study shows that map uncertainty can be very well explained by a few factors. The most important single factor is the morphology of the forest and non-forest landscape. The highest level of uncertainty occurred at or close to the forest edge, especially for type 2 error (overestimation). This is in line with our hypothesis. However, surprisingly high rates of uncertainty also occurred in forest core areas, especially for type 1 error (underestimation) where uncertainty rates for core and edge are almost equal. This seems counter-intuitive but might be related to the fact that core areas are not always well visible from ground perspective (tall trees). As expected, more inaccuracy occurred at higher elevations while against our expectations terrain slope did not play a role in the model. While one the one hand steep slopes are less accessible and therefore more difficult to map, on the other hand such slopes might be better visible from opposite slopes which makes mapping more convenient. This is a factor which is impossible to assess quantitatively, because the surveyors exact positions during the mapping process is not known. There is probably some potential to improve the model by including more complex topographical parameters such as terrain roughness. Distance to roads and visibility significantly contributed to explain type 2 errors (overestimation of

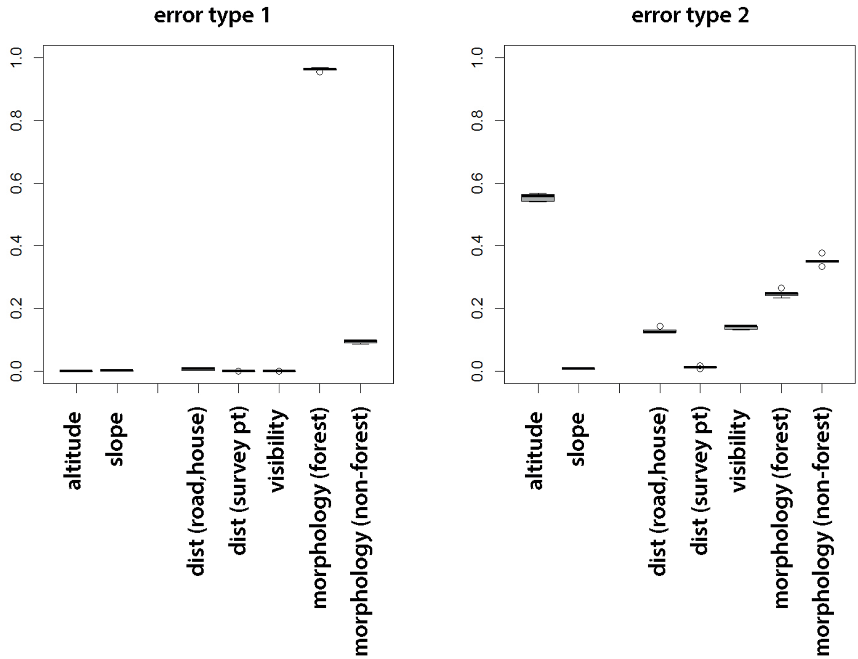

Figure 5. Variable importance for both error types

forest cover). It has to be taken into account that the two indicators are only very rough approximations to anticipate the surveyor's behaviour in the landscape. More precise information may possibly be gained by studying – if available – the surveyors notebooks.

There are different ways of applying uncertainty information when analysing historical forest cover and its dynamics. The most simple way is to choose the appropriate method and scale of analysis. For example, it does not make sense to study small scale changes based on highly uncertain forest cover information. As a simple rule, the scale of analysis should not go below the highest positional uncertainty found in the data. A good way of including uncertainty information in a qualitative way is to define areas of different quality. This can be conducted by assigning labels expressing the likelihood that the land cover information is correct at given site. Grossinger *et al.* (2007) for example used the labels definite, probable and possible in order to define a decreasing reliability of the data. Another approach to deal with uncertainty in historical map information is to compute fuzzy membership classes – in the case of forest cover forest and non-forest classes

(Leyk, Zimmermann 2007). However, for this calculation information on the nature and reasons for uncertainty is needed. The analysis of trajectories is a valuable and simple approach to assess the reliability of time series of land cover based on map information (Kaim *et al.* 2014). For time series it is also useful to check the congruence with earlier maps as, particularly in less accessible regions, information might be simply adopted from earlier maps (especially if no major change became obvious to the surveyor).

This case study confirms the importance of using multiple source types – a well-established principle in historical ecological research (Egan, Howell 2001). The reliability of one specific source (in our case historical maps) may be challenged by cross-comparing it with other independent source types (in our case orthophotos and terrestrial photos). In this context, terrestrial photos are a particularly valuable source for ground truthing remote sensing data (Kolecka *et al.* 2015).

We conclude with highlighting the importance to define and quantify as accurate as possible the uncertainty of historical map information. Knowledge on the accuracy of such information supports the creditability of the products extracted from historical maps (in this case forest cover) and all possible further analysis related to these products (e.g., estimation on carbon dynamics in forests or biodiversity assessments).

Acknowledgements. Research carried out within the FORECOM project (*Forest cover changes in mountainous regions – drivers, trajectories and implications*, PSRP 008/2010), supported by a grant from Switzerland through the Swiss contribution to the enlarged European Union. We would like to thank Dominik Kaim and Krzysztof Ostafin for interesting discussions on the topic and valuable inputs for this paper.

References

Allouche O., Tsoar A., Kadmon R. 2006, *Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS)*, Journal of Applied Ecology, 43, 1223–1232.

Boesch R., Ginzler C., 2013, *Automatische Berechnung einer nationalen Gehölzmaske*, Geomatik Schweiz, 111, 518–520.

Bozzini C., Conedera M., Krebs P., 2012, *A new monoplotting tool to extract georeferenced vector data and orthorectified raster data from oblique non-metric photographs*, International Journal of Heritage in the Digital Era, 1, 499–519.

DeFries R., Hansen A.J., Newton A., Hansen M.C., 2005, *Isolation of protected areas in tropical forests over the last twenty years*, Ecological Applications, 15, 19–26.

Egan D., Howell E.A., 2001, *The historical ecology handbook*, Island Press, New York.

Gimmi U., Lachat T., Bürgi M., 2011, *Reconstructing the collapse of wetland networks in the Swiss lowlands 1850–2000*, Landscape Ecology, 26, 1071–1083.

Figure 4. Example of a result of morphological spatial pattern analysis

wkładka _4+4_str 1 _ art Gimmi i in _ między strony 16 i 17

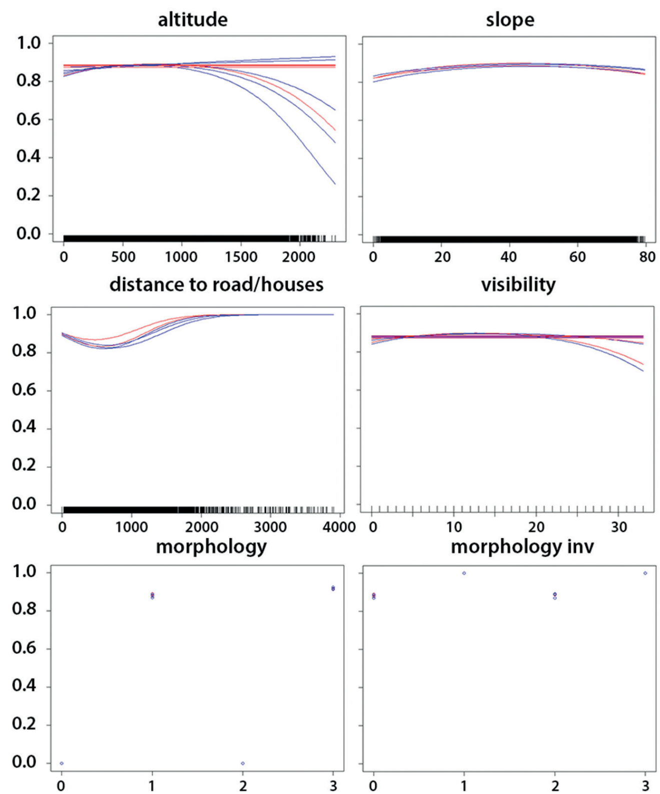

Fig. 6. Response curve of variables for error type 1.
For morphology: 0 – outside; 1 – core; 2 – gaps; 3 – edge

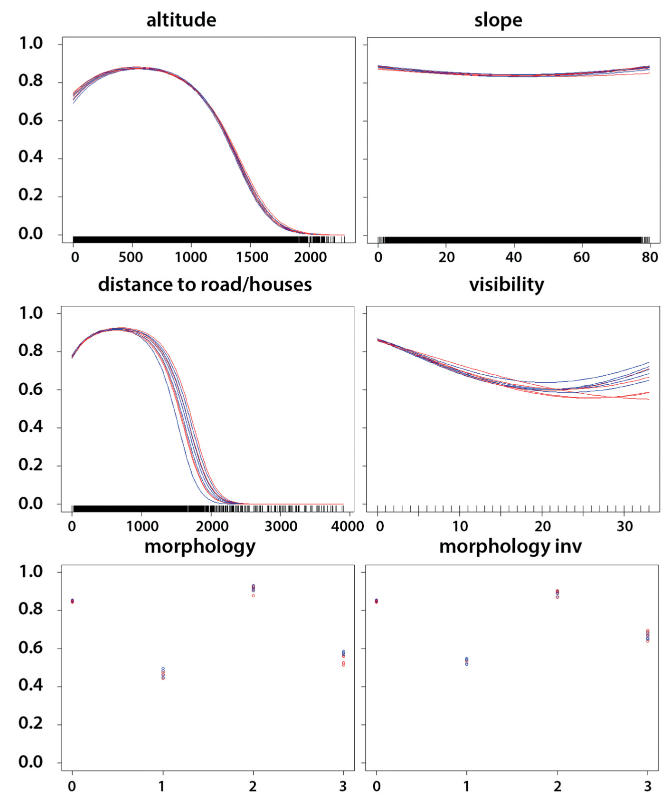

Figure 7: Response curve of variables for error type 2.
For morphology: 0 – outside; 1 – core; 2 – gaps; 3 – edge

Ginzler C., Brändli U-B., Hägeli M., 2011, *Waldflächenentwicklung der letzten 120 Jahre in der Schweiz*, Schweizerische Zeitschrift für Forstwesen, 162, 337–343.

Graae B.J., Sunde P.B., Fritzboger B., 2003, *Vegetation and soil differences in ancient opposed to new forests*, Forest Ecology and Management, 177, 179–190.

Grossinger R.M., Striplen C.J., Askevold R.A., Brewster E., Beller E.E., 2007, *Historical landscape ecology of an urbanized California valley: Wetlands and woodlands in the Santa Clara valley*, Landscape Ecology, 22, 103–120.

Gugerli D., Speich D., 2002, *Topographien der Nation. Politik, kartographische Ordnung und Landschaft im 19. Jahrhundert*, Chronos Verlag, Zürich.

Hermy M., Honnay O., Firbank L., Grashof-Bokdam C., Lawesson J.E., 1999, *An ecological comparison between ancient and other forest plant species of Europe, and the implications for forest conservation*, Biological Conservation, 91, 9–22.

Herrault P-A., Larrieu L., Cordier S., Gimmi U., Lachat T., Ouin A., Sarthou J-P., Sheeren D., 2016, *Combined effects of area, connectivity, history and structural heterogeneity of woodlands on the species richness of hoverflies (Diptera: Syrphidae)*, Landscape Ecology, 31, 877–893.

Kaim D., Kozak J., Kolecka N., Ziółkowska E., Ostafin K., Ostapowicz K., Gimmi U., Munteanu C., Radeloff V.C., 2016, *Broad scale forest cover reconstruction from historical topographic maps*, Applied Geography, 67, 39–48.

Kaim D., Kozak J., Ostafin K., Dobosz M., Ostapowicz K., Kolecka N., Gimmi U., 2014, *Uncertainty in historical land-use reconstructions with topographic maps*, Quaestiones Geographicae, 33, 55–63.

Kolecka N., Kozak J., Kaim D., Psomas A., 2015, *Mapping secondary forest succession on abandoned agricultural land with LiDAR point Clouds and terrestrial photography*, Remote Sensing, 7, 8300–8322.

Kozak J., 2003, *Forest cover changes in the Western Carpathians in the past 180 years: A case study on the Orawa region in Poland*, Mountain Research and Development, 23, 369–375.

Leyk S., Zimmermann N.E., 2004, *A predictive uncertainty model for field-based survey maps using generalized linear models*, Lecture Notes in Computer Science, 3234, 191–205.

Leyk S., Zimmermann N.E., 2007, *Improving land change detection based on uncertain survey maps using fuzzy sets*, Landscape Ecology, 22, 257–272.

Leyk S., Boesch R., Weibel R., 2005, *A conceptual framework for uncertainty investigation in map-based land cover change modelling*, Transactions in GIS, 9, 291–322.

Mather A.S., 1992, *The forest transition*, Area, 24, 367–379.

McCullagh P., Nelder J.A., 1989, *Generalized Linear Models*, Monographs on Statistics and Applied Probability 37, Chapman & Hall, Washington DC.

Munteanu C., Kuemmerle T., Keuler N.S., Müller D., Balázs P., Dobosz M., Griffiths P., Halada L., Kaim D., Király G., Konkoly-Gyuró E., Kozak J., Lieskovsky J., Ostafin K., Ostapowicz K., Sandra O., Radeloff V.C., 2015, *Legacies of 19th century land use shape contemporary forest cover*, Global Environmental Change, 34, 83–94.

Rudel T.K., Coomes O.T., Moran E., Achard F., Angelsen A., Xu J., Lambin E., 2005, *Forest transitions: Toward a global understanding of land use change*, Global Environmental Change, 15, 23–31.

Skalos J., Engstova B., Trpakova I., Santruckova M., Podrazsky V., 2012, *Long-term changes in forest cover 1780–2007 in central Bohemia, Czech Republic*, European Journal of Forest Research, 131, 871–884.

Skalos J., Weber M., Lipsky Z., Trpakova I., Santruckova M., Uhlirova L., Kukla P., 2011, *Using old military survey maps and orthophotograph maps to analyse long-term land-cover changes – case study (Czech Republic)*, Applied Geography, 31, 426–438.

Vogt P., Riiters K.H., Estreguil C., Kozak J., Wade T.G., Wickham J.D., 2007, *Mapping spatial patterns with morphological image processing*, Landscape Ecology, 22, 171–177.

Wulf M., Sommer M., Schmidt R., 2010, *Forest cover changes in the Prignitz region (NE Germany) between 1790 and 1960 in relation to soils and other driving factors*, Landscape Ecology, 25, 299–313.

Urs Gimmi
Landscape Dynamics
Swiss Federal Institute for Forest
Snow and Landscape Research
Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland
e-mail: urs.gimmi@wsl.ch

Christian Ginzler
Landscape Dynamics
Swiss Federal Institute for Forest
Snow and Landscape Research
Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

Matthias Müller
Landscape Dynamics
Swiss Federal Institute for Forest
Snow and Landscape Research
Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

Achilleas Psomas
Landscape Dynamics
Swiss Federal Institute for Forest
Snow and Landscape Research
Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland