

ASSESSING THE AGREEMENT BETWEEN EO-BASED SEMI-AUTOMATED LANDSLIDE MAPS WITH FUZZY MANUAL LANDSLIDE DELINEATION

F. Albrecht ^{a,*}, D. Hölbling ^a, B. Friedl ^a

^a Department of Geoinformatics – Z_GIS, University of Salzburg, Austria – (florian.albrecht, daniel.hoelbling, barbara.friedl)@sbg.ac.at

ISPRS ICWG III/IVb

KEY WORDS: Landslides, Fuzzy Delineation, Agreement, Accuracy, Object-based Image Analysis, Remote Sensing, Uncertainty

ABSTRACT:

Landslide mapping benefits from the ever increasing availability of Earth Observation (EO) data resulting from programmes like the Copernicus Sentinel missions and improved infrastructure for data access. However, there arises the need for improved automated landslide information extraction processes from EO data while the dominant method is still manual delineation. Object-based image analysis (OBIA) provides the means for the fast and efficient extraction of landslide information. To prove its quality, automated results are often compared to manually delineated landslide maps. Although there is awareness of the uncertainties inherent in manual delineations, there is a lack of understanding how they affect the levels of agreement in a direct comparison of OBIA-derived landslide maps and manually derived landslide maps. In order to provide an improved reference, we present a fuzzy approach for the manual delineation of landslides on optical satellite images, thereby making the inherent uncertainties of the delineation explicit. The fuzzy manual delineation and the OBIA classification are compared by accuracy metrics accepted in the remote sensing community. We have tested this approach for high resolution (HR) satellite images of three large landslides in Austria and Italy. We were able to show that the deviation of the OBIA result from the manual delineation can mainly be attributed to the uncertainty inherent in the manual delineation process, a relevant issue for the design of validation processes for OBIA-derived landslide maps.

1. INTRODUCTION

Landslide mapping benefits from the ever increasing availability of Earth Observation (EO) data resulting from programmes like the Landsat programme of NASA and the Sentinel missions of the European Copernicus programme. There is also a continuously improving infrastructure for data access. NASA provides its Landsat data for example via the Landsatlook portal (<https://landsatlook.usgs.gov/>) operated by the United States Geological Survey (USGS). For Sentinel, the access is provided by the official Sentinel Data Access Infrastructure for different user categories (e.g. <https://scihub.copernicus.eu/>), by the ESA member states, by international partner organisations and further commercial providers.

In addition to the high resolution (HR) optical satellite imagery mentioned above (Landsat, Sentinel-2), very high resolution (VHR) optical satellite imagery, e.g. WorldView, QuickBird, Pléiades, are used for landslide detection and inventory preparation (Scaioni et al., 2014; Singhroy, 2005, van Westen et al., 2008). Manual visual image interpretation is useful for mapping shallow landslides causing significant changes on the surface that are well perceivable in satellite images by their colour contrast, shape, size, pattern and texture (Morgan, 2010), or in case of stereographic imagery elements of the third dimension (Scaioni et al., 2014). Drawbacks of manual interpretations are that they are resource- and time-consuming and highly dependent on the experience and skills of the interpreter (Galli et al., 2008; Guzzetti et al., 2012; Hölbling et al., 2015). As shown by various studies in other fields (Albrecht,

2010, Albrecht et al., 2010, Kohli et al., 2016; Van Coillie et al., 2014), human visual interpretation tasks generally underlie variability over individuals and time, i.e. no interpreter will achieve the very same result twice and no interpreter will achieve the very same result as his peer. Still, manual landslide interpretation represents the current standard for landslide detection and inventory preparation. Despite its value, the lack of repeatability renders manual interpretations somewhat subjective.

Regarding the increased availability of EO data for landslide mapping and the limited resources to fully exploit these data by means of visual interpretation, the disadvantages of manual interpretation and their lack of repeatability becomes even more evident. Advances in computer science and machine intelligence facilitate the development of automated methods for landslide mapping (Moosavi et al., 2014; van Westen et al., 2008). However, Scaioni et al. (2014) emphasize that remote sensing techniques for landslide investigations still require further development, even if recent improvements are very promising. Both pixel-based and object-based image analysis (OBIA) approaches are available for the analysis of optical data, often used in combination with digital elevation models (DEMs) (Stumpf and Kerle, 2011). The recent trend however, is towards OBIA approaches as demonstrated by a range of landslide mapping examples (e.g. Blaschke et al., 2014; Barlow et al., 2006; Behling et al., 2014; Heleno et al., 2016; Hölbling et al., 2012, 2015, 2016, 2017; Kurtz et al., 2014; Lahousse et al., 2011; Lu et al., 2011; Martha et al., 2010, 2011, 2012; Stumpf and Kerle, 2011; Rau et al., 2014). OBIA provides a set of image processing tools and algorithms that can be jointly applied for semi-automatically delineating and classifying

* Corresponding author

landslides on EO data. The tools apply algorithms that rely on the concepts of image segmentation and classification. For landslide recognition based on EO data, landslides or landslide elements are aggregated into homogeneous objects and are then classified according to specific characteristics. These characteristics represent the spectral, spatial, textural and contextual properties of the image objects (Barlow et al., 2006; Martha et al., 2010; Moosavi et al., 2014).

OBIA approaches definitely show potential for the adoption in operational processes for landslide inventory preparation. But in order to prove their validity for this purpose, OBIA results have to be comparable to the results of the current state-of-the-art method of manual visual image interpretation. Consequently, manually delineated landslide maps often are used as a reference for judging the quality of OBIA results using common accuracy assessment techniques from remote sensing (Congalton and Green, 2009; Foody, 2002), applied for OBIA-derived landslide maps by Hölbling et al. (2016). A recent study by Hölbling et al. (2017) compares manual mapping results to semi-automated object-based mapping results and shows that both methods can produce results of similar quality, whereby the achieved accuracy values can vary widely between study areas.

The study of Hölbling et al. (2017) also argues that the term ‘agreement’, instead of ‘accuracy’, is better suited for the comparison of semi-automated object-based mapping results to manual mapping results. A traditional accuracy assessment requires reference data with significantly higher accuracy (Lang et al., 2010). This is a quality level that manual mapping performed on the same image data cannot fulfil due to its inherent uncertainties originating from the variability in the interpretation process. This does not disqualify the accuracy metrics that measure the agreement between the datasets. But it acknowledges that a certain amount of disagreement between the datasets cannot be avoided. Although there is awareness of the uncertainties inherent in manual delineations, there is a lack of understanding how they affect the levels of agreement in a direct comparison of OBIA-derived landslide maps and manually derived landslide maps. If the impact of the uncertainty in the manual delineations is better understood, the validity of the agreement values as a quality measure for the OBIA-derived landslide maps can be better assessed.

This study analyses the agreement between OBIA-derived landslide maps and a fuzzy manual delineation of landslides that have been prepared on the basis of the same optical satellite images. The fuzzy approach makes the inherent uncertainties in the interpreter’s choice of the manual delineation explicit. Albrecht (2010) shows that the drawn delineations in the interpretations specifically differ (1) in areas where a spectral gradient indicates a transition from one class to another, (2) in areas where an interpreter chooses to generalize instead of following every detail provided by the image, and (3) in areas where the image information remains uncertain about the applied class definition. With a fuzzy manual delineation approach for mapping landslides, the variations of type 1 and to some degree of type 2 can be addressed by identifying a spatial bandwidth of variation. The fuzzy delineation is compared to an OBIA-derived landslide map with accuracy metrics commonly used in remote sensing. The calculated metrics will enable to judge if the OBIA-derived landslide map falls within a similar range of uncertainty as has been made explicit with the fuzzy manual delineation, thereby estimating the agreement between the compared datasets. The result is discussed in the context of

the utilization of manual delineation for assessing the quality of OBIA-derived landslide maps.

2. MATERIAL AND METHODS

2.1 Study areas and data

Landslides in two study areas in the Alps were investigated: (1) the Fürweg landslides at the Haunsberg, Salzburg, Austria, and landslides near the villages of (2) Badia and (3) Corvara that are located in the Gader Valley, South Tyrol, Italy (Figure 1). As a basis for the landslide mapping, Landsat 7 and Sentinel-2 data were available.

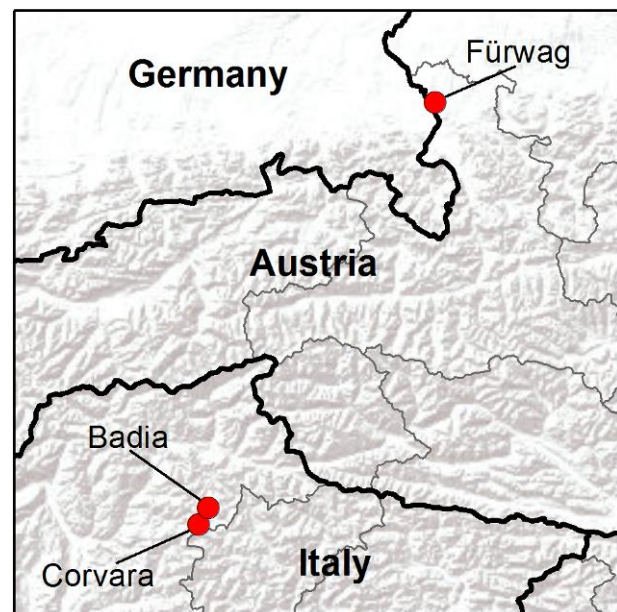


Figure 1: Overview of study areas. Sources: ESRI, USGS, NOAA.

The ‘Fürweg Landslide’ at the Haunsberg, about 10 km north of the city of Salzburg, Austria, is situated in an area with remnants of mudslides that occurred in previous times. The most recent sliding processes took place between 1999 and 2003 when major infrastructure facilities at the valley floor were endangered (Fiebiger, 2002). The geology in this pre-alpine zone is characterized by sandstones, shales and the Flysch, formed by the Zementmergelserie (marl cement series) which are very prone to mass movements (Embleton-Hamann, 2007).

The Gader Valley in South Tyrol, Italy, exhibits particular geological and geomorphological conditions that make it prone to landslides (Larcher et al., 2010), e.g. containing the area of the Corvara landslide with a size of more than 2.5 km² and a volume of 300 million m³ (Corsini et al., 2005). We focus on two recent landslides in its wider neighbourhood: in December 2012, long and intensive rainfall caused a large landslide in Sottrù in the municipality of Badia that destroyed several buildings and endangered infrastructure; in April 2014, a large landslide happened at the border of Corvara to Colfosco.

For the study area Fürweg, a Landsat 7 satellite image from 28/07/2002 with 15 m resolution in the panchromatic band and 30 m resolution in the multispectral bands was used. Pre-processing included a pansharpening and the orthorectification

of the image to match the coordinate system WGS1984 UTM 33N with sub-pixel geolocation accuracy. A digital elevation model (DEM) with 10 m resolution, freely available from the geodata portal of Austria (www.geoland.at), was used as ancillary data. It originated from the Austrian airborne laser scanning (ALS) campaigns between 2006 and 2013.

For the study areas of Badia and Corvara, a Sentinel-2 satellite image from 27/08/2016 was available. The four multispectral bands with 10 m resolution (blue, green, red, nir) were used, whereas the other bands in 20 m and 60 m resolution were disregarded. The image was transformed to WGS1984 UTM 32N with sub-pixel geolocation accuracy. A 5m DEM was available from the geodata portal of the Autonomous Province of Bolzano (<http://geoportale.buergernetz.bz.it/default.asp>).

2.2 Object-based landslide mapping

For each of the study areas, a semi-automated object-based mapping approach was developed in eCognition (Trimble Geospatial) in order to identify the landslides in the satellite images (cf. Hölbling et al., 2017). The approaches followed similar workflows, but segmentation parameters and classification rules had to be slightly adapted to each test case since the spectral differences in the datasets limited the transferability. Image objects were created with multiresolution segmentation and selected object features were used in the classification ruleset for discriminating the landslide areas from other areas. They included the Normalized Difference Vegetation Index (NDVI), the slope information and further spectral and spatial parameters (e.g. spectral difference to neighbouring objects, length/width ratio, shape index). An exploratory analysis of the image segments' properties determined the selection of thresholds. Table 1 summarizes the used segmentation and classification parameters for each study area.

Study Area	Segmentation parameters	Main classification parameters
Fürwag	SP: 25 S: 0.1 C: 0.5	Mean diff. to neighbours (NDVI) < 0 Mean slope > 10 Length/Width > 4
Badia	SP: 150 S: 0.5 C: 0.5	Mean NDVI < 0.4 or < 0.5 (depending on fine adjustment in combination with other features) Mean slope > 10
Corvara	SP: 150 S: 0.3 C: 0.8	Mean NDVI < 0.4 Mean slope > 8 Length/Width > 1.6 Shape index < 3.6

Table 1. Segmentation parameters (SP – Scale Parameter, S – Shape criterion, C – Compactness criterion) and classification parameters for the ruleset applied to each study area (cf. Hölbling et al. (2017)).

2.3 Fuzzy manual landslide mapping

For the fuzzy manual delineation, landslides were mapped by drawing polygons with the editing tools available in ArcGIS

(ESRI). The digitizing scale was set to 1:20.000. The satellite images as the main source of information were displayed by the combination of the bands red, green and blue, i.e. in true-colour. Other band combinations and derived features (slope, hillshade) of the DEM were used as auxiliary data. Due to the low spatial resolution of the satellite images relative to the scale, each pixel of the satellite image was visible in the display, because nearest neighbour resampling from satellite image pixel to screen pixel was used to display the image. Additionally, filtered versions of the satellite images were overlaid, using a resampling of the satellite image pixels to the screen pixels with cubic convolution and setting the transparency to 50%. Thereby, in addition to the pixel boundaries, the spectral gradient from one pixel centre to the next became visible. This allowed targeting the uncertainty in the delineation of the border between areas that are part of a landslide and areas that are not part of a landslide.

We used a fuzzy membership (FM) approach to produce a manually derived map with three different delineations of the landslides. The delineations separate the area of interest into four membership classes according to the likelihood that the area actually belongs to a landslide. The four classes are:

- **LS_{cert}**: FM=1.0; the likelihood that the area belongs to a landslide is 1.0 and not lower, i.e. the interpreter is certain that the delineation only includes landslide areas;
- **LS_{uncert}**: FM>0.5; the landslide likelihood is higher than (or equal to) 0.5 and lower than 1.0;
- **Non-LS_{uncert}**: FM<0.5; the landslide likelihood is higher than 0.0 to lower than 0.5;
- **Non-LS_{cert}**: FM=0.0; the landslide likelihood is 0.0, i.e. no landslide is present.

Therefore, the delineation between the classes **LS_{uncert}** and **Non-LS_{uncert}** targets the threshold where the likelihood that the location belongs to a landslide is equal to the likelihood that the location is not a landslide. With this approach, mainly the spectral uncertainty in the boundary definitions of image objects was targeted that is inherent to the spatial resolution of the satellite image and related blurring effects. After digitizing the three delineations, their consistency among each other was checked at a scale of 1:10 000 to confirm that the lines did not intersect.

2.4 Comparison method

For comparing the semi-automated mapping results from OBIA to the FM manual landslide delineation, the traditional thematic accuracy assessment approach was used (Congalton and Green, 2009; Foody, 2002; Pontius and Millones, 2011). The comparison method has successfully been applied to landslide classifications by Hölbling et al. (2016). Essentially, the traditional thematic accuracy assessment uses the proportion of correctly classified areas. The classification that shall be tested is compared to a reference dataset for identifying areas where the compared datasets agree and where they disagree (Figure 2). Two types of error can occur from the perspective of the landslide class: the error of omission occurs in areas where the reference suggests a landslide that is not represented in the classification; the error of commission occurs in areas where the classification suggests a landslide that is not represented in the reference. The user's accuracy (UA) value expresses the correctly classified area as a share of the classification area that includes errors of commission. The producer's accuracy (PA)

expresses the correctly classified area as a share of the reference area that includes errors of omission.

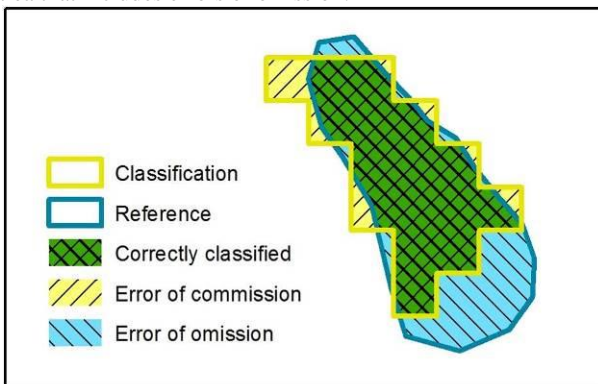


Figure 2: Error types in a remote sensing image classification.

When the classification and the reference have only two classes, i.e. 'landslide' and its complement 'non-landslide', the resulting 2-by-2 contingency matrix is relatively simple. In the case of this study, the OBIA-derived landslide classification has two classes whereas the fuzzy manual landslide delineation has four classes. Consequently, we arrive at a 2-by-4 contingency matrix

that enables the analysis of error at different thresholds of uncertainty. UA and PA were calculated for each acceptance level of uncertainty, i.e. $FM=1.0$, $FM>0.5$ and $FM>0.0$. E.g. for $FM>0.5$, $UA_{FM>0.5}$ and $PA_{FM>0.5}$ are calculated from the summarized matrix values of all the classes that fulfil this uncertainty level, i.e. LS_{cert} and LS_{uncert} . In addition, some metrics compared the total area size of the LS_{OBIA} class to each of the summarized areas for each acceptance level of uncertainty from the manually derived reference. These were put into relation to the $FM>0.5$ area as well.

3. RESULTS

3.1 Object-based and fuzzy manual mapping results

The results of the semi-automated object-based mapping and the fuzzy manual mapping for the study areas of Fürwag (Figure 3), Badia (Figure 4) and Corvara (Figure 5) are shown in the following.

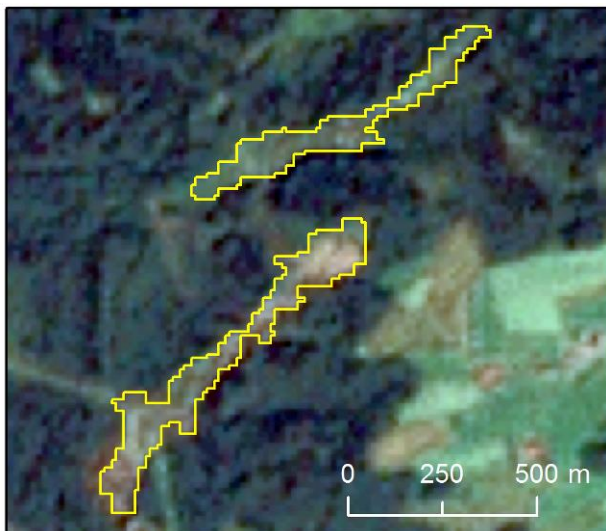


Figure 3: Semi-automated mapping of landslides achieved with object-based image analysis (in yellow colour; left) and fuzzy manual mapping of landslides (in red colour; right) in the study area of Fürwag based on the Landsat 7 image.

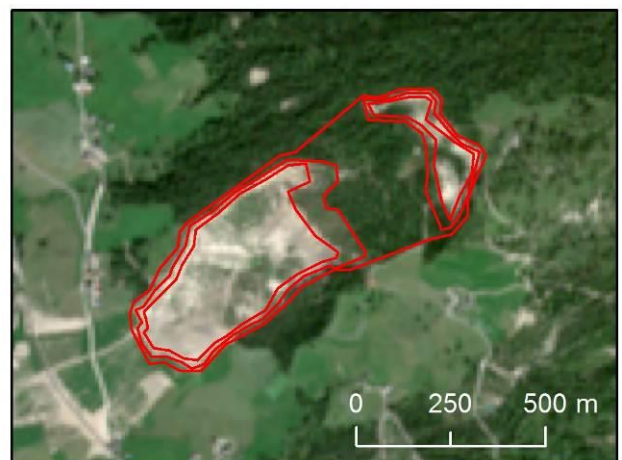
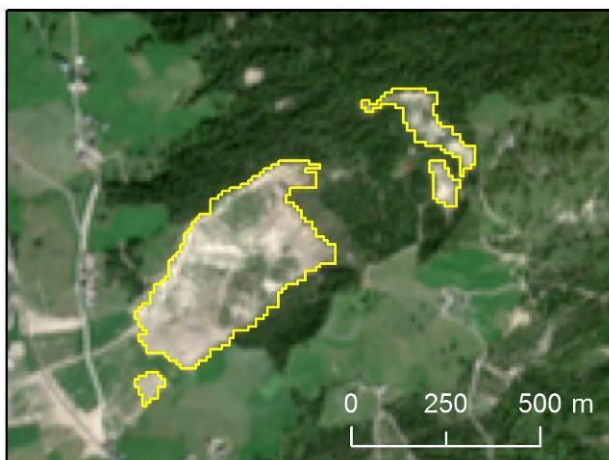


Figure 4: Semi-automated mapping of landslides achieved with object-based image analysis (in yellow colour; left) and fuzzy manual mapping of landslides (in red colour; right) in the study area of Badia based on the Sentinel-2 image.

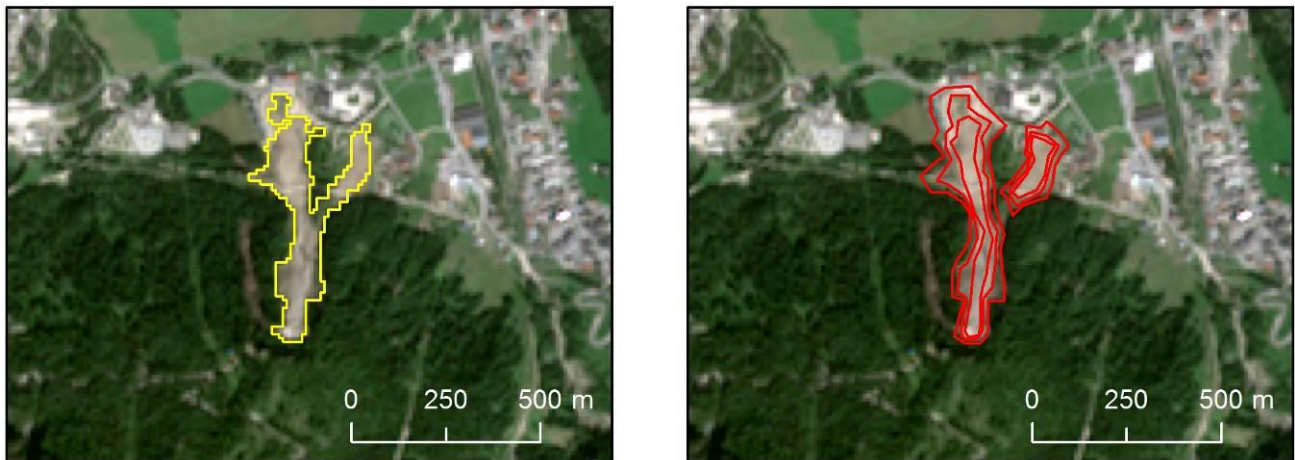


Figure 5: Semi-automated mapping of landslides achieved with object-based image analysis (in yellow colour; left) and fuzzy manual mapping of landslides (in red colour; right) in the study area of Corvara based on the Sentinel-2 image.

The visual comparison between the results of the two methods shows a good coincidence in all the study areas. However, on the manually delineated lowest likelihood level, the landslide polygon near Badia (Figure 4) includes a vegetated area as part of the landslide that has not been indicated by the object-based result, since there is no spectral indication of a landslide. The human interpreter is able to logically infer that with some likelihood an area can be assigned to the landslide class that is not immediately perceivable as landslide (Hölbling et al., 2017). On the other hand, the study area of Corvara (Figure 5) showed a branching landslide in the object-based result that the fuzzy manual mapping has identified as two separate landslides. This may be an error in the fuzzy manual mapping or an error in the semi-automated result. Obviously, spectral image information at the 10 m resolution of Sentinel-2 can be ambiguous when extracting landslides.

3.2 Comparison results

The calculated metrics compare the OBIA mapping result to the fuzzy manual landslide mapping result. Contingency matrices were provided for each of the three study areas Fürwag (Table 2), Badia (Table 4), and Corvara (Table 6). The metrics derived from the contingency matrices were compiled in the corresponding tables, i.e. Table 3 for Fürwag, Table 5 for Badia and Table 7 for Corvara. The latter set of tables presents the metrics by the different levels of accepted uncertainty. FM=1.0 presents metrics for the contingency matrix column LS_{cert}, FM>0.5 presents the metrics for the two summarized contingency matrix columns of LS_{cert} and LS_{uncert}, and FM>0.0 presents the metrics for the three summarized contingency matrix columns of LS_{cert}, LS_{uncert} and Non-LS_{uncert}. Thereby, the difference between the OBIA result and the fuzzy manual landslide mapping reference dataset is made explicit.

OBIA classes	Fuzzy Manual Delineation				Sum area
	LS _{cert}	LS _{uncert}	Non-LS _{uncert}	Non-LS _{cert}	
LS _{OBIA} (km ²)	0.084	0.038	0.019	0.007	0.148
Non-LS _{OBIA} (km ²)	0.003	0.017	0.043	---	---
Sum area (km ²)	0.087	0.055	0.062	---	

Table 2: Contingency matrix for study area Fürwag.

Accuracy metric	Accepted level of uncertainty		
	FM>0.0	FM>0.5	FM=1.0
OBIA Mapping area (km ²)	0.148		
Manual Mapping area within threshold (MM; km ²)	0.204	0.142	0.087
Difference OBIA to MM (% of MM)	-27.8	4.2	69.2
Difference MM to FM>0.5 (% of FM>0.5)	44.3	0.0	-38.4
Overlap area (km ²)	0.141	0.122	0.084
Producer's accuracy PA (%)	68.8	86.0	96.1
User's accuracy UA (%)	95.2	82.5	56.8

Table 3: Comparison values for study area Fürwag.

OBIA classes	Fuzzy Manual Delineation				Sum area
	LS _{cert}	LS _{uncert}	Non-LS _{uncert}	Non-LS _{cert}	
LS _{OBIA} (km ²)	0.139	0.028	0.008	0.005	0.180
Non-LS _{OBIA} (km ²)	0.007	0.034	0.114	---	---
Sum area (km ²)	0.146	0.062	0.122	---	

Table 4: Contingency matrix for study area Badia.

Accuracy metric	Accepted level of uncertainty		
	FM>0.0	FM>0.5	FM=1.0
OBIA Mapping area (km ²)	0.180		
Manual Mapping area within threshold (MM; km ²)	0.330	0.208	0.146
Difference OBIA to MM (% of MM)	-45.5	-13.4	22.9
Difference MM to FM 0.5 (% of FM 0.5)	58.8	0.0	-29.5
Overlap area (km ²)	0.175	0.167	0.139
Producer's accuracy PA (%)	52.9	80.4	94.7
User's accuracy UA (%)	97.1	92.8	77.1

Table 5: Comparison values for study area Badia.

OBIA classes	Fuzzy Manual Delineation				Sum area
	LS _{cert}	LS _{uncert}	Non-LS _{uncert}	Non-LS _{cert}	
LS _{OBIA} (km ²)	0.027	0.026	0.014	0.003	0.070
Non-LS _{OBIA} (km ²)	0.001	0.004	0.029	---	---
Sum area (km ²)	0.028	0.031	0.043	---	---

Table 6: Contingency matrix for study area Corvara.

Accuracy metric	Accepted level of uncertainty		
	FM>0.0	FM>0.5	FM=1.0
OBIA Mapping area (km ²)	0.070		
Manual Mapping area within threshold (MM; km ²)	0.102	0.059	0.028
Difference OBIA to MM (% of MM)	-31.6	17.7	147.9
Difference MM to FM 0.5 (% of FM 0.5)	72.0	0.0	-52.5
Overlap area (km ²)	0.067	0.053	0.027
Producer's accuracy PA (%)	65.6	89.8	96.9
User's accuracy UA (%)	95.9	76.3	39.1

Table 7: Comparison values for study area Corvara.

When comparing the results by the total mapped area, the OBIA results correspond best with the threshold of FM>0.5, with values between -13.4% and 17.7%. In relation to FM=1.0 the OBIA results overestimate the area, in relation to FM>0.0 they underestimate the area. The total areas of the fuzzy manual mapping results vary strongly, where FM>0.0 is at least twice as large as FM=1.0 and, in case of the study area of Corvara, almost 4 times as large. From FM>0.5 they deviate up to 72%. This defines the range of uncertainty captured by the fuzzy manual mapping approach within which errors of omission and errors of commission have to be expected.

For the accuracy metrics of UA and PA, Figure 6 shows their behaviour over the different accepted uncertainty thresholds.

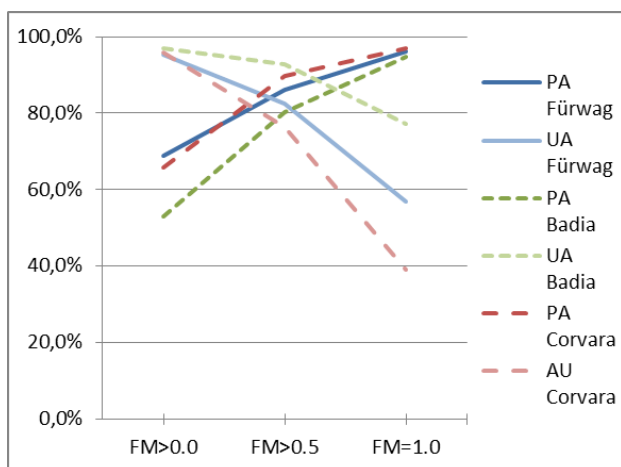


Figure 6: User's and producer's accuracies for each study area in relation to the different accepted uncertainty thresholds.

If the accuracy of the summarized landslide area is relevant to the subsequent analysis, then the errors of omission and errors of commission should cancel each other out. This is the case when the values of PA and UA are as close to each other as possible. Figure 6 shows that for the analysed case studies, the threshold FM>0.5 best fulfils this criterion. Other cases of subsequent analysis may favour one error type over the other. The delineation of the threshold FM>0.0 includes areas towards

the outside of the landslide whereas the other two thresholds already count the extent of the OBIA result's as errors of commission. FM>0.0 consequently achieves very high UA values above 95% and relatively low PA values below 69%. In reverse, FM=1.0 excludes areas towards the inside of the landslide that the other two delineations already count as errors of omission. Accordingly, PA reaches values above 95% and UA values below 77%, for Corvara even only 39%.

4. DISCUSSION

The objective of this study was to better understand the uncertainty inherent to landslide maps. Ideally, the results would lead to a better estimation of acceptable levels of disagreement between results from semi-automated landslide mapping approaches and results from manual delineation of landslides. Therefore, we performed tests in three study areas where an OBIA classification and a fuzzy manual delineation were produced, followed by a comparison with accuracy assessment metrics. The comparison results were interpreted for identifying the sources of error that underlie the uncertainties inherent to EO-based landslide mapping and for identifying the different handling of uncertainty in the manual approach and the OBIA approach.

We were able to show that the deviation of the OBIA results from the fuzzy manual delineations can be attributed to the uncertainty inherent to manual landslide delineation to a large degree. The uncertainty is introduced by the limited spatial resolution of satellite images where digitized lines can only reach a spatial accuracy that is lower than the pixel size, decreased by blurring effects. The limited spatial resolution also limits the mapping scale used for manual mapping. We tried to increase this by utilizing a smoothed display of the pixels of the satellite image. However, this approach may only help for comparing mapping results that originate from the same image. If landslide maps from a different origin were used for comparison, the positional error from the image orthorectification process would need to be accounted for, as well. The OBIA delineations are bound to the stepped pixel border originating from the segmentation process. Therefore, deviations from the manual mapping that affect accuracy values are inevitable. The above described uncertainties affect both, the results from semi-automated OBIA mapping and from manual mapping. Consequently, a manual delineation cannot be significantly more accurate than an OBIA mapping result if the same image is used as data base. The uncertainties made explicit by the applied fuzzy manual mapping are reflected in the calculated accuracy metrics. They show how the measured errors fall within the range of uncertainty. Therefore, the term 'accuracy' is somewhat misleading and following Hölbling et al. (2017), we suggest using the term 'agreement' instead.

Nevertheless, some differences occurred between the OBIA and the fuzzy manual mapping that cause a type of disagreement that comparison metrics should report. We found areas, specifically in the study area of Badia and Corvara, where the manual mapping and the OBIA result are outside of the accepted range of uncertainty and the occurrence of a misclassification, either in the manual mapping or the OBIA result, is very likely. Concerning manual mapping, the underlying process of visual perception for the interpretation of images is complex and causes uncertainty. According to Albrecht (2010), the uncertainty in the interpretation process of

drawing delineations of image objects is related to the spectral gradient between neighbouring class objects, to the generalisation applied by the interpreter and to thematic uncertainty resulting from the limited resolution of the image. An interpreter's best opportunity to address the resulting uncertainty is to make it explicit. This may happen by fuzzy manual delineation, as presented in this study, or by documenting the quality of crisp landslide delineations. For their manual delineations of landslides, Hölbling et al. (2017) document the achievable geometric accuracy in relation to the image resolution and the interpreter's subjective evaluation of the achieved completeness of the manual mapping results. Concerning potential uncertainty in the OBIA result, there are fuzzy class assignment approaches available that assign segments to different classes with fuzzy membership values (Feizizadeh et al., 2017). Also the thematic uncertainty in assigning classes has been captured by an accuracy assessment that distinguishes different weights of error for gully erosion mapping (D'Oleire-Oltmanns et al. 2014). However, the mentioned research addresses more the thematic uncertainty and less the uncertainty in the spatial delineation of the resulting objects, an issue that can be addressed in future research.

Even with a fuzzy manual mapping approach, we cannot be certain that the approach is free of error. An accuracy assessment with a reference of much higher accuracy and from an independent source would be required. Essentially, accuracy values can only be calculated, if a reference exists that is more accurate by at least one magnitude of the measurement scale. This would be achievable if we had a reference on hand, that was delineated from an aerial photograph or ideally a reference that was created by a survey with in-situ measurements on the ground providing the highest accuracy that can be achieved, in due consideration that uncertainty is still present.

Finally, the quality assessment of semi-automated landslide mapping shall be applicable in more practical settings and may be implemented in the context of an EO-based web service for landslide mapping (Hölbling et al. 2017). In this respect, a different type of reference than provided by the fuzzy manual mapping is required, as it is too time-consuming to prepare. The presented study contributes to a better judgement of the uncertainty in manual mapping. Visual interpretation results can be regarded as a valid reference, when taking into consideration the concept of 'agreement' rather than 'accuracy'. As suggested by the Quality Assurance Framework for Earth Observation (QA4EO; <http://QA4EO.org>), the users of EO-based landslide maps within the landslide community will be the ones to suggest quality indicators that help them to decide the appropriateness of the landslide maps for their purpose.

ACKNOWLEDGEMENTS

This research has been supported by the Austrian Research Promotion Agency FFG in the Austrian Space Applications Program (ASAP 11) through the project Land@Slide (contract n° 847970). The authors thank Clemens Eisank, Filippo Vecchiotti, Arben Kociu and Elisabeth Weinke for fruitful discussions and the anonymous reviewer for his constructive critical input.

REFERENCES

- Albrecht, F., 2010. Uncertainty in image interpretation as reference for accuracy assessment in object-based image analysis. In: Tate, N.J., Fischer, P.F. (Eds.), *Accuracy 2010. Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. MPG Books Group, Leicester, United Kingdom, pp. 13-16.
- Albrecht, F., Lang, S., Hölbling, D. 2010. Spatial accuracy assessment of object boundaries for object-based image analysis. In: *GEOBIA 2010 - Geographic Object-Based Image Analysis*, Ghent, Belgium, 29 June - 2 July 2010, Addink, E., Van Coillie, F.M.B., (Eds.), ISPRS Vol.No. XXXVIII-4/C7, Archives ISSN No 1682-1777.
- Barlow, J., Franklin, S., Martin, Y., 2006. High spatial resolution satellite imagery, DEM derivatives, and image segmentation for the detection of mass wasting processes. *Photogrammetric Engineering and Remote Sensing*, 72(6), pp. 687-692.
- Behling, R., Roessner, S., Kaufmann, H., Kleinschmit, B., 2014. Automated Spatiotemporal Landslide Mapping over Large Areas Using RapidEye Time Series Data. *Remote Sensing*, 6(9), pp. 8026-8055.
- Blaschke, T., Feizizadeh, B., Hölbling, D., 2014. Object-Based Image Analysis and Digital Terrain Analysis for Locating Landslides in the Urmia Lake Basin, Iran. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(12), pp. 4806-4817.
- Congalton, R.G., Green, K., 2009. Assessing the accuracy of remotely sensed data - principles and practices, 2nd ed., CRC Press, Boca Raton, USA, p. 183.
- Corsini, A., Pasuto, A., Soldati, M., Zannoni, A., 2005. Field monitoring of the Corvara landslide (Dolomites, Italy) and its relevance for hazard assessment. *Geomorphology*, 66(1-4), pp. 149-165.
- D'Oleire-Oltmanns, S., Marzloff, I., Tiede, D., Blaschke, T., 2014. Detection of Gully-Affected Areas by Applying Object-Based Image Analysis (OBIA) in the Region of Taroudannt, Morocco. *Remote Sensing* 6, 8287.
- Embleton-Hamann, C., 2007. Geomorphological hazards in Austria. In: *Geomorphology for the Future*. Embleton-Hamann, C., Stötter, J., (Eds.), Innsbruck University Press, Innsbruck, Austria, pp. 33-56.
- Feizizadeh, B., Blaschke, T., Tiede, D., Moghaddam, M.H.R., 2017. Evaluating fuzzy operators of an object-based image analysis for detecting landslides and their changes. *Geomorphology* 293, Part A, 240-254.
- Fiebigler, G., 2002. Die Rutschung Fürwag im Norden von Salzburg / Österreich. Prozess und Maßnahmen. In: *Proceedings of the International Congress Interpraevent*, Matsumoto, Japan, pp. 629-639.
- Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), pp. 185-201.

- Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., Reichenbach, P., 2008. Comparing landslide inventory maps. *Geomorphology*, 94(3-4), pp. 268-289.
- Guzzetti, F., Cardinali, M., Reichenbach, P., Carrara A., 2000. Comparing Landslide Maps: A Case Study in the Upper Tiber River Basin, Central Italy. *Environmental Management*, 25(3), pp. 247-263.
- Heleno, S., Matias, M., Pina, P., Sousa, A.J., 2016. Semiautomated object-based classification of rain-induced landslides with VHR multispectral images on Madeira Island. *Natural Hazards and Earth System Sciences*, 16, pp. 1035-1048.
- Hölbling, D., Füreder, P., Antolini, F., Cigna, F., Casagli, N., Lang, S., 2012. A Semi-Automated Object-Based Approach for Landslide Detection Validated by Persistent Scatterer Interferometry Measures and Landslide Inventories. *Remote Sensing*, 4(5), pp. 1310-1336.
- Hölbling, D., Friedl, B., Eisank, C., 2015. An object-based approach for semi-automated landslide change detection and attribution of changes to landslide classes in northern Taiwan. *Earth Science Informatics*, 8(2), pp. 327-335.
- Hölbling, D., Betts, H., Spiekermann, R., Phillips, C., 2016. Identifying Spatio-Temporal Landslide Hotspots on North Island, New Zealand, by Analyzing Historical and Recent Aerial Photography. *Geosciences*, 6, 48.
- Hölbling, D., Eisank, C., Albrecht, F., Vecchiotti, F., Friedl, B., Weinke, E., Kociu, A., 2017. Comparing Manual and Semi-Automated Landslide Mapping Based on Optical Satellite Images from Different Sensors. *Geosciences*, 7(2), 37.
- Hölbling, D., Weinke, E., Albrecht, F., Eisank, C., Vecchiotti, F., Friedl, B., Osberger, A., Kociu, A., 2017. A web service for landslide mapping based on Earth Observation data, In: *Proceedings of ReSyLAB 2017 - 3rd Regional Symposium on Landslides in the Adriatic-Balkan Region*, Ljubljana, Slovenia, pages pending.
- Kohli, D., Stein, A., Sliuzas, R., 2016. Uncertainty analysis for image interpretations of urban slums. *Computers, Environment and Urban Systems*, 60, pp. 37-49.
- Kurtz, C., Stumpf, A., Malet, J.-P., Gançarski, P., Puissant, A., Passat, N., 2014. Hierarchical extraction of landslides from multiresolution remotely sensed optical images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, pp. 122-136.
- Lahousse, T., Chang, K.-T., Lin, Y., 2011. Landslide mapping with multiscale object-based image analysis—A case study in the Baichi watershed, Taiwan. *Natural Hazards and Earth System Sciences*, 11, pp. 2715-2726.
- Lang, S., Albrecht, F., Kienberger, S., Tiede, D., 2010. Object validity for operational tasks in a policy context. *Journal of Spatial Science*, 55, 9-22. doi: 10.1080/14498596.2010.487639
- Larcher, V., Notarnicola, C., Piacentini, D., Pinter, T., Schneiderbauer, S., Soldati, M., Strada, C., 2010. Analyse flachgründiger Massenbewegungen mittels Verwendung zweier statistischer Methoden im Gadertal (Südtirol). In: *Proceedings of COGeo*, Salzburg, Austria, 11.06.2010, Marschallinger, R., Wanker, W., Zobl, F., (Eds.), pp. 1-11.
- Lu, P., Stumpf, A., Kerle, N., Casagli, N., 2011. Object-oriented change detection for landslide rapid mapping. *IEEE Geoscience and Remote Sensing Letters*, 8(4), pp. 701–705.
- Martha, T.R., Kerle, N., van Westen, C.J., Jetten, V., Vinod Kumar, K., 2012. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, pp. 105-119.
- Martha, T.R., Kerle, N., van Westen, C.J., Jetten, V., Kumar, K.V., 2011. Segment Optimization and Data-Driven Thresholding for Knowledge-Based Landslide Detection by Object-Based Image Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 49(12), pp. 4928-4943.
- Martha, T.R., Kerle, N., Jetten, V., van Westen, C.J., Kumar, K.V., 2010. Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods. *Geomorphology*, 116(1-2), pp. 24-36.
- Morgan, J.L., Gergel, S.H., Coops, N.C., 2010. Aerial Photography: A Rapidly Evolving Tool for Ecological Management. *BioScience*, 60(1), pp. 47-59.
- Moosavi, V., Talebi, A., Shirmohammadi, B., 2014. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by Taguchi method. *Geomorphology*, 204, pp. 646-656.
- Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15), pp. 4407-4429.
- Rau, J.-Y., Jhan, J.-P., Rau, R.-J., 2014. Semiautomatic object-oriented landslide recognition scheme from multisensor optical imagery and DEM. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2), pp. 1336-1349.
- Scaioni, M., Longoni, L., Melillo, V., Papini, M., 2014. Remote Sensing for Landslide Investigations: An Overview of Recent Achievements and Perspectives. *Remote Sensing*, 6(10), pp. 9600-9652.
- Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. *Remote Sensing of Environment*, 115(10), pp. 2564-2577.
- Singhroy, V., 2005. Remote sensing of landslides. In: *Landslide Hazard and Risk*, Glade, T., Anderson, M.G., Crozier, M.J., (Eds.), Wiley & Sons, West Sussex, UK, pp. 469-492.
- Van Coillie, F.M.B., Gardin, S., Anseel, F., Duyck, W., Verbeke, L.P.C., De Wulf, R.R., 2014. Variability of operator performance in remote-sensing image interpretation: the importance of human and external factors. *International Journal of Remote Sensing*, 35(2), pp. 754-778.
- Van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology*, 102(3-4), pp. 112-131.