Department of 3D Mapping
BSF Swissphoto GmbH
Mittelstrasse 7
12529 Schönefeld, Germany

Schönefeld, 30th August 2017

Dear Editor,

We are pleased to submit our replies to the Reviewer's comments of manuscript NHESS-2017-120 entitled "Regional snow-avalanche detection using object-based image analysis of near-infrared aerial imagery" article.

We would like to thank two reviewers for their constructive comments, and appreciate their insightful analysis of our study. We considered all comments to improve the quality of the article. Following the suggestions, we rephrased potentially misleading statements, and added more descriptions of the sensitivity analyses and the thresholds that we have selected in our approach. We also elaborated more on the limits and problems tied to the classification protocol. We further expanded the limitations and the conceptual model of the OBIA approach in the discussion section. Finally, we corrected the figures according to a reviewer's request, and checked the text for spelling mistakes and clarity of language.

For the convenience, we address the reviewers' comments as outlined below:
- Reviewer's comments (**bold**),
- Our answers (*italic*), and the changes made in the text (*italic*).

Yours sincerely,

Karolina Korzeniowska
(on behalf of co-authors)

General comments:
The paper by K. Korzeniowska and co-workers describes a method for automatic avalanche detection based on object-based image analysis (OBIA) of near-infrared (NIR) aerial imagery. The paper starts with an introduction of the dangers and risks related to snow avalanches with some detailed information on human fatalities and damages to infrastructures (mostly focused on the case of Switzerland), and the need for further developments concerning the large-scale yet precise mapping of avalanche activity. Section 2 provides the reader with an overview of the existing studies on automatic detection of avalanche activity based on image processing. A very short section (current section 3) describes the study area and data, and already includes some information on the methods. Section 4 is the detailed section on the methods. Section 5 shows the main results with a focus on the following points: the accuracy of the OBIA workflow proposed, including the influence of variables chosen -brightness, NDVI, NDWI, SD_NDWI- on accuracy, on the one hand, and the results on (i) topographic factors that influence avalanche activity, (ii) density of avalanches and (iii) classification in release, flow and run-out zone of detected avalanches on the other hand. Section 6 is a discussion on the advantages and drawbacks (points to be improved) regarding the method proposed. Finally, the last section concludes by mostly giving a summary of the main points discussed in section 6.

The paper presents an interesting method for detection of avalanche activity, which is based on object-based image analysis (OBIA) of near-infrared (NIR) aerial imagery. Given that NIR aerial images of good resolution are available, this method has the potential to cover large-scale mountainous areas including very remote areas. Furthermore, a statistical approach allows to distinguish between release, flow and run-out zones of the detected avalanches. One tricky yet important issue remains the distinction between single and multiple events. The OBIA workflow and the results presented in the paper are worth to be published in NHESS. However, the paper in its current form is not ready for publication. Some substantial revision is needed at least for two main reasons. First, I found that the sensitivity analysis of the method and outcomes to varying the different thresholds is lacking. The authors should pay attention to including such a sensitivity analysis when possible, and/or argue more on their choice regarding the thresholds of a number of parameters: see my specific comments below. Second, the authors should make an effort to improve the editing/structure of their manuscript that is sometimes quite hard to follow: short section versus another much longer section, announcement of outlines needed in the main introduction and the long section on "methods", etc. (see my comments regarding Editing/Structure of the paper and technical corrections).

We thank the Reviewer for this positive assessment. We took into account all the suggestions provided by the Reviewer to improve the quality of our manuscript. Specifically, we focused in our revisions on improving the description of sensitivity analysis. We had long discussions within the author team about the best structure, and aimed for a more accessible structure of the workflow. Please see below for our specific responses:

Specific comments:
- section 1 (intro), page 2 (lines 20-21): Saying "..., focussing mainly on geographic coordinates, but rarely on any detailed information about their extent or area." Appears too strong to me. I think the authors should qualify that statement. Traditional methods, such as photo-interpretation, the use of testimonies, photographs, etc, gives crucial information, and merging that information into one single platform (see for instance the paper by Irstea group, by Bourova et al. CRST 2016, as well as some references therein) is already an important and efficient step. Could you please revise this part of the text?

We agree with the Reviewer that this sentence might be misleading. We agree that photo-interpretation allows us to obtain detailed information about the shape and the size of avalanches. Our intention here was to stress that at the national level (e.g. in Switzerland) information about snow avalanches is usually collected as geographic point coordinates; this information can be obtained without any remote sensing data. However, coordinates are added to snow-avalanche databases mostly only in case of damage or accident. The Reviewer noted that more detailed information about avalanche geometry can be collected for areas where a detailed remote sensing data are available. However, using aerial images in the winter season is problematic, as we pointed out on page 2, lines 30-33), so that collecting detailed information at the national level remains challenging.

We rephrased the sentence in our manuscript to avoid misunderstandings of its interpretation as follow, page 2, lines 20-21:

"To date, experts map most avalanches manually, and at the national level mainly the geographic coordinates of avalanches; rarely does this mapping involve detailed information about the avalanche geometry (Bühler et al., 2009)."

- section 2, page 3 (lines 23-25): Please could you argue a little bit on this 35 deg threshold? I believe such a limit should depend on the snow type and then influence the results of your automatic detection approach. Please could you comment on this point?

We point out that this slope threshold is not part of our study, so that we do not elaborate on that. Figure 7 shows that the avalanches that we mapped manually and used as reference data occur also on steeper slopes. Bühler et al. (2009) noted that using a slope threshold may misclassify some detachment areas of avalanches; therefore we excluded any assumptions regarding slope in our OBIA approach.

- section 2, page 4 (line 25): again, this threshold of 35 deg needs much more discussion. This angle should depend on the snow type. Either you show a sensitivity analysis of your detection to that threshold, or you give more physical arguments on this choice.

Please see our reply above: we did not use any slope threshold in our study. We were referring to its use in previous studies.

- section 4.1.1, page 6 (line 4): why this threshold of 6.25 square meters (the exponent 2 for the segment area is missing)? Could you please argue to make your choice less arbitrary? Did you conduct any sensitivity analysis of your method to varying that threshold? If yes, could you please include a thorough discussion on that analysis in the manuscript?

The exponent 2 for segment is fine, we checked this in an online version of the article and it appears fine there. We selected the 6.25 square meters threshold by a trial-and-error approach in the eCognition software by visual checking the results for various thresholds. Applying this threshold allowed us to obtain the highest correctness in classifying those parts of avalanches that did not meet the brightness threshold of snow avalanches. We made this clearer on page 6, lines 12-14:

"To select the best size of segments we ran a sensitivity analysis using a trial-and-error approach in the eCognition software, visually checking the results for various thresholds."

- section 4.1.2, page 6 (lines 13-16): the sentence is not clear to me, please revise. Could you please show a couple of examples of the cross-comparison between segment values and their visual representation in an image?

We agree that this sentence may not be fully clear. What we wanted to say is that inside some segments that we classified as avalanches we noticed small areas of smooth snow not representing avalanche debris and thus in need of reclassification. We rephrased this sentence to make it clearer, and added some example of cross-comparison between segment values and their visual representation in Figure 4. See page 6, lines 12-16:

"We thus included many pixels as parts of avalanche deposits that in the previous step had escaped being classified as 'rough snow' because of SD_{NDWI} values. We set a maximum area of segment to be reclassified as 'rough snow' to avoid including large, but smooth, areas inside avalanches that are not avalanche debris."

- end of section 4.1.2: could you please argue on this threshold of 62.2 square meters? Again: did you conduct any sensitivity analysis? Also I would suggest you show a square of that size on one figure (among figs a,b,c,d shown in Fig. 4 for instance). It would be useful for the reader to materialize the physical size of that threshold onto the map, compared to avalanche extensions that are detected.

We selected the threshold of 62.2 square meters after running a sensitivity analyses in the eCognition software, and visually checking the results when changing the thresholds. Following the Reviewer's suggestion, we added squares in Figure 4 to show the size of this threshold with respect to the size of classified avalanches. Please see the changes in the text on page 6, lines 29-31:

"We chose this area threshold based on visual checks and a trial-and-error sensitivity analysis in the eCognition software."

- section 4.1.4, page 7 (line 4): could you please explain why you are using/choosing those values for both brightness and NDVI thresholds? I would suggest you to show a sensitivity analysis to varying those thresholds. This seems to me a crucial point if you would like your method to be possibly extended to much larger scales or other mountainous areas.

We used these thresholds because they fitted well in separating our data into those representing snow and those that do not represent snow. The sensitivity analyses for Brightness and NDVI are already presented in Figure 6. We added additional sentence in this matter, please see page 7, lines 12-14:

"The brightness thresholds we derived from a sensitivity analysis, in analogy to previous steps (Fig. 6)."

- section 4.1.5, page 7 (lines 24-25 in brackets): why those thresholds? Could you show a sensitivity analysis to varying the thresholds? Brightness threshold is 2500 here, while it was 3000 a little earlier in the text (see previous comment too).

Please see our reply above; the results of the sensitivity analyses are shown in Figure 6. The thresholds in this sentence (and this stage of classification) are different because some parts of snow avalanches were a bit darker or a little vegetated, so applying the previous threshold assumptions did not allow their correct assignment to an avalanche class. Therefore, we lowered these thresholds based on the sensitivity analyses in Figure 6 to allow for correct classification. We added a sentence on page 8, lines 4-6:

"The threshold selected for brightness is lower than in previous steps, because some parts of snow avalanches were a little darker; hence using the same threshold would cause unwanted misclassifications."

- section 5, page 8 (lines 17-22): this part discusses the problems/limits of the method. The reader would like to know how those problems/limits are sensitive to the choice of the different thresholds (see the specific comments above). Could you please strengthen the discussion on this point?

We added more detail about the limits and problems that occur during the classification process to the discussion section, please see the changes on page 12, lines 27-35:

"The most difficult to classify were old avalanches where the bare ground cropped out or where vegetation occurred in the path of the avalanche. These avalanches did not meet the assumptions in our OBIA protocol and could not be classified correctly because they were not rough enough, or were too vegetated or dark due to thin snow cover. Here a solution may be using different thresholds for old and fresh avalanches to optimise our OBIA approach. We tested different thresholds for the input layers and different neighbourhood assumptions in order to include these avalanches; however, this resulted in more false positives, so the cost of correct classification of these avalanches was higher than the benefit. We therefore decided to stay with the same workflow and thresholding shown in Figure 4 for the whole study area."

- section 5.3, page 9 (line 12): is the range 20-40 deg compatible with the threshold of 35 deg discussed in section 2?

No, they are not compatible. Our results show that avalanches can occur also in slopes steeper than 35 degree. Coming back to our reply to a similar comment, we note that we did not incorporate any specific slope thresholds. We added a sentence in this matter, please see page 9, lines 14-17:

"Our results show that avalanches were also detected on slopes steeper than 35°, thus highlighting the limits of detection methods using arbitrary slope thresholds (Bühler et al., 2009; Vickers et al., 2016)."

Editing/Structure of the paper and Technical corrections (typings errors, etc.):
- Abstract (line 19): "... at three 4-km areas..." The exponent 2 for areas in square kilometres per square is missing?

We have checked this issue in the whole manuscript carefully and could not reproduce these errors.

- end of section 1 (intro), page 3 (lines 14-15): again, the exponent 2 is missing here. Please fix it.

Please see our response above.

- section 1: would be nice to terminate the section by announcing the detailed structure of the paper, with a short summary of each section (including explicit numbered reference to each section).

We did discuss the possibility of an additional short summary within the author team, but agreed that it is not necessary and would make the paper unduly longer.

- section 2, page 3 (lines 23-25): the construction of the sentence is a bit weird... The threshold of 35 deg is an assumption (a model input that stems from the DEM data) but not an outcome of the RAMMS model. Please revise the sentence.

We restructured this sentence to avoid misunderstandings. Please see page 3, lines 28-30:

"They used the numerical simulation tool RAMMS (Rapid Mass Movement Simulation; Christen et al., 2010) to identify possible avalanches, and excluded slopes >35° from the runout calculation, assuming that these slopes could not accumulate snow-avalanche debris."

- section 2, page 3 (line 27): "nadir?". Please fix this.

Yes, the word "nadir" is fine.

- section 3, page 5 (line 10): again, the exponent 2 is missing... PLEASE CHECK CAREFULLY the entire text regarding this issue, which is present in many parts of the manuscript.

As we mentioned above, we think that this issue did not come up on our side. In the entire manuscript submitted all exponents are fine.

- section 3 is very very short. Given the fact that it already includes some information on the methods, I would suggest that the authors include section 3 as a sub-section of section 4 "Methods".

We agree that this section is very short, but would maintain its structure focusing on some basics of the study area.

- introduction of section 4: this part needs careful revision. Would be nice to put here an outline with explicit references to the sub-sections that follow (by using the numbering), as well as explicit reference to key figure 4. In its current form, I must say that section 4 is very difficult to read/follow.

We added to the text the numbering of subchapters to make the text easier to read, and also the numbering of subchapters to Figure 4. Please see the changes in the text on page 5, lines 14-20:

"We introduce an automatic method for mapping release zones, tracks, and runout zones of avalanches using NIR 0.25-m aerial images (Section 4.1). We compare the automatic classification with manually digitised reference data and estimated the accuracy of detecting snow avalanches with confusion matrices (Sections 4.2, 5.1). We also investigated the topographical conditions in which most mapped avalanches occurred (Section 5.3), using two approaches for visualising avalanche density (Section 5.4). We also propose a probability approach to representing release and runout zones of avalanches and the automatic classification of snow avalanche parts (Section 5.5.)."

- end of section 4.1.2, page 6 (line 19): typo... should be "then" instead of "than"

OK, we changed "than" to "then".

- section 5.4: I would suggest to replace "inset 1, Fig. 8" by "inset on Fig. 8a" and "inset 2 on Figure 8" by "inset on Fig 8b".

We changed this accordingly.

O. Jaquet (Referee)

olivier.jaquet@in2earth.com

This paper presents an automatic object-based image analysis approach for the detection of snow avalanches using specific remote sensing data. The approach is explained in details and some statistical tests are performed for the evaluation of results uncertainty in relation to a reference data set. Following additional explanations/ comments would be valuable for the reader:

I.   A description of the conceptual model associated to the OBIA approach, it would provide some insight in relation to uncertainties and limitations; since only 61% of the total avalanche area was correctly identified...

We elaborated on the limitations and the conceptual model of the OBIA approach in the discussion section of our manuscript. The 61% of the total avalanche area correctly classified comes mostly from the observations that many avalanches classified were not fresh, so their properties differed from those of fresh avalanches. We addressed this matter on page 12, lines 22-27:

"Our method correctly identified 61% of the total avalanche area in the study area, mainly because of heterogeneous avalanche debris with dark or smooth patches, and because of differing deposit ages, and hence differing surface roughness."

II.   Mathematical details of the selected probability approach; to my understanding an uniform distribution is selected in relation to the elevation; the selection of such simple statistical model should be justified in relation to point I.

Unfortunately, we do not fully understand the point of this comment. We would be grateful for providing more detail information in this matter.

III.   A reference for confusion matrices

The reference for confusion matrices which we used are provided in the manuscript on page 8, lines 19-22:

"We used them to estimate several performance metrics, including Type I, Type II, and total errors (Sithole and Vosselman, 2004), overall, user's, and producer's accuracies (Congalton, 1991), Cohen's kappa (Cohen, 1960), and F-Score"

IV.   The impact of snow thickness on the results

We did not analyse the impact of snow thickness on the results, mainly limited by the low accuracy of available snow-thickness maps. However, we added a sentence to the discussion section, please see page 13, lines 11-14:

"Detailed enough maps of snow thickness may help to improve the classification accuracy, and perhaps enable separating release zones, tracks, and runout zones from the surrounding snow cover."

V.   The potential use of additional data (radar, seismic, ...) to improve the results

We are not sure if we understand this comment correctly. In general, we think that the radar data and other seismic data provide further opportunities for detecting snow avalanches, and we mentioned

that in the introduction chapter by citing Eckerstorfer and Malnes (2015). We think that, combining the results obtained from aerial images and radar data may increase the accuracy of hazards maps. However, we are a bit sceptic by combining aerial images and radar data in classifying snow avalanches with OBIA, especially because the collection of these data would have to be synchronized. We added a sentence to the discussion section, please see page 13, lines 22-23:

"Future tests involving diverse areas across the World and different types of data (e.g. UAV, radar, and seismic) are desirable."

Regional snow-avalanche detection using object-based image analysis of near-infrared aerial imagery

Karolina Korzeniowska[1, 2], Yves Bühler[3], Mauro Marty[4], Oliver Korup[2],

[1]3D Mapping, BSF Swissphoto GmbH, Schönefeld, 12529, Germany
5 [2]Geohazards Research Group, University of Potsdam, Potsdam, 14476, Germany
[3]WSL Institute for Snow and Avalanche Research SLF, Davos, 7260, Switzerland
[4]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, 8903, ~~Swizterland~~Switzerland

Correspondence to: Karolina Korzeniowska (karolina.korzeniowska@bsf-swissphoto.com)

10 **Abstract.** Snow avalanches are destructive ~~natural hazards~~mass movements in mountain regions that continue to claim lives, and cause infrastructural damage and traffic detours. Given that avalanches often occur in remote and poorly-accessible steep terrain, their detection and mapping is extensive and time consuming. Nonetheless, systematic avalanche detection over large areas could help to generate more complete and up-to-date inventories (cadastres) necessary for validating avalanche forecasting and hazard mapping. In this study, we ~~have~~ focused on automatically detecting avalanches and classifying them
15 into release zones, tracks, and runout zones based on 0.25-m near-infrared (NIR) ADS80-SH92 aerial imagery using an object-based image analysis (OBIA) approach. Our algorithm takes into account the brightness, the normalised difference vegetation index (NDVI), ~~and~~ the normalised difference water index (NDWI) and its standard deviation (SD_{NDWI}) ~~in order~~ to distinguish avalanches from other land-surface elements. Using normalised parameters allows ~~readily~~ applying this method across large areas. We trained the method by analysing the properties of snow avalanches at three 4-km² areas near Davos,
20 Switzerland. We compared the results with manually-mapped avalanche polygons, and obtained a user's accuracy of >0.9 and a Cohen's kappa of 0.79 – 0.85. Testing the method for a larger area of 226.3 km², we estimated producer's and user's accuracies of 0.61 and 0.78, respectively, with a Cohen's kappa of 0.67. Detected avalanches that overlapped with reference data by >80% occurred randomly throughout the testing area, showing that our method avoids overfitting. Our method ~~shows~~has potential ~~in~~for large-scale avalanche mapping, although further investigations into other regions are desirable to
25 verify the ~~stability~~robustness of our selected thresholds and the transferability of the method.

1 Introduction

Snow avalanches are frequent and destructive mountain hazards, particularly during the winter and spring months. They are fast mass movements controlled by weather conditions, snowpack, and ~~the~~ topography ~~of the terrain~~ (Schweizer et al., 2003; Castebrunet et al., 2012). Avalanches can cause loss of lives, disrupt infrastructure, and ~~affect~~ bury buildings (Bründl et al.,
30 2004; McClung and Schaerer, 2006; Eckerstorfer and Malnes, 2015).

Despite numerous efforts aimed at reducing the risk ~~posed by~~from avalanches, most fatalities in Europe occur during sporting activities, caused by avalanches triggered by the victims themselves (Techel et al., 2015). Past research indicates that poor decision-making and forecasting are the main causes of deadly avalanche accidents (Techel et al., 2015; McClung, 2016). Techel et al. (2015) stated that most destructive events occur on days when the snow avalanche risk is very critical

5   and the snowpack layer is weak. In total, 4,750 people lost their lives in the European Alps between 1970 and 2015 (Techel et al., 2016); in the past two decades, avalanches ~~in the Swiss Alps alone~~ have killed 461 people in the Swiss Alps alone (Fig. 1). Most fatal accidents have occurred in the cantons of Valais and Grison, which are the two largest in Switzerland and contain some of the ~~highest~~ most-elevated areas in the Swiss Alps. Since 1946, avalanches in Switzerland have had the highest share of victims (37%) ~~in comparison~~compared to other natural hazards, including lightning (16%), floods (12%),

10 windstorms (10%), rockfalls (8%), and landslides (7%; Badoux et al., 2016).

Avalanches killed a total of 36 people in Switzerland in the fatal winter of 1998/99 (Fig. 1). Between 27 January and 25 February, 17 people died in villages and on roads, and material losses surpassed 600 million Swiss Francs. This catastrophic winter spurred an initiative aimed at improving avalanche safety and reducing concomitant losses (Wilhelm et al., 1999; Bründl et al., 2004). This initiative included establishing an information system for exchanging data between the WSL

15 Institute for Snow and Avalanche Research SLF and local authorities; the development of hazard maps showing zones with high avalanche susceptibility (Bründl et al., 2004); and an increase in the artificial release of avalanches to decrease hazard levels. In this context, the need for documenting avalanches also increased, to allow for learning from past accidents.

This initiative also showed clearly that regional-scale mapping of avalanches and identifying potential release zones is not only desirable, but ~~also~~ essential for producing avalanche cadastre maps ~~to be used~~ for quality-checking hazard mapping and

20 forecasting (Bühler et al., 2015). To date, experts map most avalanches manually, and at the national level mainly the geographic coordinates of avalanches; rarely does this mapping involve detailed information about the avalanche geometry (Bühler et al., 2009). ~~To date, experts (Bühler et al., 2009) map most avalanches manually, focussing at the country level mainly on avalanches geographic coordinates, but rarely on any detailed information about their extent or area.~~ ~~Moreover~~Hence, avalanche inventories are biased toward damaging events or those reported from accessible terrain, so that~~.~~

25 ~~Hence,~~ avalanches remain notoriously underreported over larger regions. To more broadly collect information concerning avalanches, non-expert observers in Switzerland are now able to report sightings via an app (http://www.slf.ch/lawinenbulletin/rueckmeldung/index_EN), where they can enter the location and date of their observation. Optical remote sensing data, both airborne and satellite, offer coverage that is more systematic, and are therefore increasingly used to track avalanches. Satellite images allow ~~the collection of a picture for~~covering the same area

30 with a time interval equal to one satellite orbit around the Earth. Airborne images can be acquired even more often, although in the winter season it is not preferable to continue the campaign due to financial ~~aspects~~constraints, because such images, which ~~represent~~ feature mostly only snow, are not convenient for any purposes other than the assessment of the risk of snow avalanches.

Automatic methods for detecting snow avalanches are still in the developing stage (Eckerstorfer et al., 2016), and different kinds of data, such as optical and radar images, and classification approaches are used to verify their suitability to track avalanche events. This motivated us to ~~verify~~ explore the ~~usability~~ use of near infrared (NIR) aerial images and their ~~calculated~~ derivatives ~~for,~~ in mapping avalanches over a large area in Switzerland, as well as ~~verifying~~ investigating the

5   topographic conditions of their occurrence. We ~~have~~ proposed an automatic OBIA-based method for detecting avalanche runout zones, ~~as well as~~ their tracks, and release areas. We tested whether normalised indices of water and vegetation derived from aerial ADS80-SH92 images are suitable in this regard, and introduced a simple method for roughly distinguishing these zones, because knowledge concerning potential release zones critically aids ~~in~~ hazard assessments and runout models. Our motivation was to develop an algorithm widely applicable to mountain regions~~:, as~~ relying on image spectral properties alone

10 (Lato et al., 2012) may limit such portability, because objects may have a ~~similar~~ brightness similar to that of snow.

Most avalanches start on slopes with a median inclination of 39° (Schweizer and Jamieson, 2001), ~~hence~~ on slopes most difficult for skiing. We distinguish two types of avalanche: loose avalanches start from a point and gradually increase in size as they move downslope, whereas slab avalanches involve the detachment of large planar packs of snow (Fig. 2; Schweizer et al., 2003; Bagli and Schweizer, 2009). Path length defines whether an avalanche is small (10 – 100 m), medium (100 –

15 1,000 m), or large (>1,000 m; Eckerstorfer et al., 2016). Most large avalanches~~, such as that shown in Figure 2,~~ are slab avalanches because to bring down a large amount of snow, planar snow detachment is necessary (Fig~~ure~~. 2); however, new smaller loose snow avalanches may occur and overlap with the previous one, thus complicating their detection in the field. In terms of avalanche deposit area, we distinguish between large deposits (>2,000 m²), small deposits (100 – 2,000 m²), and very small deposits (< 100 m²; Bühler et al., 2009). Every avalanche has a release zone – a part where the avalanche is

20 triggered, a track, a part where the snow is transported down the slope – and a runout, or deposition, zone (Fig. 2).

2 Previous work

Most previous work devoted to mapping avalanches from optical remote sensing data has focused on delineating runout zones. The idea of using object-based image analysis (OBIA) for detecting avalanches has been used in conjunction with brightness information from aerial images and local slope data ~~taken~~ from digital elevation models (DEM), and with

25 numerical modelling (Bühler et al., 2009), whereas others have used only the spectral information of aerial and satellite images ~~to detect snow avalanches~~ (Lato et al., 2012). Bühler et al. (2009) ~~proposed an approach for~~ mapp~~ed~~ing snow-avalanche deposits from ADS40 20-cm aerial images, which they resampled to 1 m and then combined with 25-m elevation data. They used the numerical simulation tool RAMMS (Rapid Mass Movement Simulation; Christen et al., 2010) to identify ~~areas of~~ possible avalanches, ~~and used a DEM to~~ excluded slopes >35° from the runout calculation, ~~as they~~ assum~~ed~~ing that

30 these slopes could not accumulate snow-avalanche debris. They also ~~used~~ worked with spectral thresholds to exclude snow-free areas. To separate rough avalanche debris from surrounding smooth and undisturbed snow, they used the normalised difference angle index (NDAI), evaluated from nadir and backwards NIR bands. They computed the NDAI difference between neighbouring pixels with a grey-level co-occurrence matrix (GLCM), which represents the distribution of pixel

3

values at a given offset, and found that the thresholding of an entropy measure evaluated via the GLCM achieved the best separability of rough and smooth snow. Ski lifts and other objects characterised by a similar entropy were removed using OBIA. The estimated accuracy of this method in terms of the fraction of correctly-detected avalanche deposits was 94%, and the producer's accuracy was 87%. Lato et al. (2012) applied OBIA for detecting avalanche deposits from panchromatic

5   images only. They tested their algorithm with Quick Bird images in Norway, and aerial ADS40 images in Switzerland, relying on six variables (i.e. GLCM entropy, GLCM dissimilarity, brightness, contrast, similarity, and neighbour distance) in their procedure, in which segments failing to meet the OBIA assumptions were sequentially discarded. They started by eliminating dark regions from brightness data before detecting rough snow with edge contrast~~s~~. The similarity filter and density helped to remove isolated pixels and small objects, respectively. Finally, the neighbour distance helped to fill gaps

10   inside the extracted snow-avalanche deposits. The user's and producer's accuracies of ~~this~~ that classification were both >90%. Both these studies (Bühler et al., 2009; Lato et al., 2012) regarded OBIA as suitable for detecting snow-avalanche deposits, because it considers the spatial relation of the analysed segments in addition to their spectral properties.

Larsen et al. (2013) suggested an approach for optical Quick Bird imagery using directional filters evaluated based on image texture classification (Varma and Zisserman, 2004) to distinguish avalanches from ~~an~~other objects. They assumed that

15   avalanches have a linear texture pattern ~~with a linear structure on the snow that~~ ~~ec~~coincid~~ing~~es with the local hillslope aspect. Similar to other strategies, their classification took into account neighbouring pixels, while parameters such as area, area-perimeter ratio, aspect direction difference, co-occurrence mean, correlation, and entropy ~~were used to assist in~~helped to exclud~~ing~~e misclassified instances. Based on a visual comparison, the authors concluded that their classification was acceptable, allowing the detection of many fresh avalanches with a low number of false alarms (Larsen et al., 2013). They

20   pointed out, however, that some of the detected avalanches were split into parts; they therefore recommended additional processing to re-connect those fragments.

Eckerstorfer and Malnes (2015) manually detected avalanche debris based on its higher backscatter contrast, compared to the surrounding undisturbed snow cover, in Radarsat-2 Ultrafine SAR imagery. They assumed that avalanches are tongue-shaped features with high surface roughness and higher snow density than surrounding terrain. Surface roughness and snow

25   density were determined from backscatter, which increased in cases of higher surface roughness, and absorption, which increased for denser snow, respectively. They found that release zones and tracks were mostly difficult to detect. In a similar context, the automatic method ~~of~~ by Vickers et al. (2016) evaluates backscatter in 50 x 50 pixel regions of Sentinel-1A images, subsequently masking out areas with a predicted zero probability of snow-avalanche occurrence ~~of zero~~; pixels with a DEM-derived local slope of >35° were ~~removed~~ignored ~~from the occurrence of avalanches~~. From the test pixels, they

30   selected those with a backscatter difference above a specified threshold. Randomly-selected pairs of pixels gave a total dissimilarity of pixels and class representatives for a K-mean clustering with two classes, i.e. 'avalanche' or 'not avalanche'. The ~~detection rate (~~producer's accuracy~~)~~ of this algorithm was ~~60%~~0.6, and the authors highlighted its potential for avalanche monitoring despite having to mask~~ing~~ out large amounts of data. ~~Finally,~~ Bühler et al. (2016) tested an unmanned aerial vehicle (UAV) that allows for fast, repeatable, flexible, and cost-efficient measurements of snow depths in alpine

4

terrain, possibly generating digital surface models of homogenous snow surfaces (Bühler et al., 2017). Legal regulations in Switzerland and elsewhere currently limit broad coverage of UAV imagery, however.

3 Study area and data

Our study area is centred around Davos, in the Swiss canton of Grisons; the area has alpine relief, with the highest local peak
5 at Schwarzhorn (3,146 m a.s.l.; Fig. 3). Many slopes in this area exceed 28° and have dominant northeastern and southwestern aspects (Fig. 3). We used 0.25-m resolution NIR aerial images in conjunction with abundant avalanche information acquired via a ADS80-SH92 large-format Digital Pushbroom Sensor (Leica Geosystems AG, Heerbrugg, Switzerland; Bühler et al., 2009) at the end of the 2012/13 winter season. The sensor recorded information with five spectral bands: panchromatic, blue, green, red, and NIR (Bühler et al., 2009). We have data for the area for more than five time slices
10 starting from the winter of 2007/08 and continuing from 2011/12 to 2015/16. The images we used were taken at the end of the winter of 2012/13, where the highest expected snow depths were between 2,000 and 3,000 m a.s.l., and covered ~226 km².

4 Methods

We ~~have~~ introduced an automatic method for mapping release zones, tracks, and runout zones of avalanches using NIR 0.25-
15 m aerial images (Section 4.1). We compared the automatic classification with manually -digitised reference data and ~~assessed~~ estimated the accuracy of detecting snow avalanches ~~using~~ with confusion matrices (Sections 4.2, 5.1). ~~Furthermore, w~~We ~~verified~~ also investigated the topographical conditions ~~i~~on which most mapped avalanches occurred (Section 5.3), ~~and verified~~using two approaches for visualising ~~the~~ avalanche density (Section 5.4). ~~In addition, w~~We also proposed a probability approach to representing release and runout zones of avalanches and the automatic classification of
20 snow avalanche parts (Section 5.5.).

4.1 Automatic OBIA snow- avalanche classification

We implemented a multi-step OBIA approach for detecting avalanches in eCognition Developer 9.1.1 software (Fig. 4). As input for the classification, we used the green, red, and NIR bands, and computed from these the normalised difference vegetation index (NDVI = $\rho_{NIR} - \rho_{Red} / \rho_{NIR} + \rho_{Red}$; Townshend and Justice, 1986), and the normalised difference water index
25 (NDWI = $\rho_{Green} - \rho_{NIR} / \rho_{Green} + \rho_{NIR}$; McFeeters, 1996) and its standard deviation in a 5 x 5 kernel (SD$_{NDWI}$). We derived brightness as the mean of the green, red, and NIR bands to classify 'dark objects', such as rivers, rocks, and buildings. The NDVI ~~helped~~ helps to classify trees, bushes, and other types of vegetation, whereas the NDWI (SD$_{NDWI}$) detect~~s~~ed snow (~~rough snow;~~ Fig. 4). We stretched thus computed NDVI and NDWI data into an interval of [0, 255].

5

4.1.1 Classifying vegetation, dark objects, and snow

In ~~our~~ a first step, we segmented the data using a chessboard segmentation algorithm in eCognition 9.1.1 ~~software~~, assigning a standalone segment to each ~~single~~ pixel (Fig. 4). We used the pixel NDVI and brightness values of snow-free areas ~~obtained from NDVI and brightness~~ to classify 'vegetation' and 'dark objects'. We classified vegetation as having positive NDVI values (corresponding to >127 in the stretched range), and dark objects as those with a brightness of <4,000 (Fig. 4). Because pixels on the border of vegetation and dark objects have mixed values, we assigned them to a separate 'buffer' class by reclassifying every pixel that shared a border with either 'vegetation' or 'dark objects' (existence of 'vegetation' or existence of 'dark objects'; Fig. 4), and then excluded this 'buffer' class from further analysis. We merged all segments classified as 'vegetation' and 'dark objects' and ~~assigned~~ treated all segment areas <6.25 m² (<100 pixels; Fig. 4) as ~~being~~ too small to buffer or divert an avalanche, and included these segments as potential avalanche areas ~~where a snow avalanche could occur~~. All the size thresholds ~~used~~ in our OBIA workflow were set to the resolution of the data ~~that were used~~ and the size of the ~~analysed~~ mapped avalanches. To select the best size of segments we ~~performed~~ ran ~~the~~ a sensitivity analysis using a trial-and-error approach in the ~~the~~ eCognition software, ~~by visual~~ly ~~interpretation of~~checking the results ~~when~~for ~~changing~~various ~~the~~ thresholds. Similarly, we classified pixels with a positive NDWI (>127 for stretched data) as snow and pixels with an additional roughness contrast (SD_{NDWI} >1) as 'rough snow', representing avalanches.

4.1.2 Removing small objects that do not represent avalanches

We then merged all snow pixels and reclassified 'rough snow' segments <12.5 m² (< 200 pixels, Fig. 4). We thus included many pixels as parts of avalanche deposits that in the previous step had escaped being classified as 'rough snow' ~~in the previous step~~ because of too low SD_{NDWI} values ~~that were too low~~. We ~~applied an assumption concerning the~~set a maximum area of segment ~~that can~~to be reclassified ~~into as~~the 'rough snow' ~~class~~ to avoid ~~the~~ inclu~~sion of~~ding large, but smooth, areas inside avalanches ~~which~~that ~~in fact do not represent an~~are not avalanche debris. After comparing the segment values with their visual representation ~~in an image~~, we observed that ~~larger~~ smooth areas ~~that are~~ inside ~~the~~ of avalanche debris ~~represent were small ascents that were omitted by the avalanche~~areas of flow separation, ~~and should therefore not be assigned as a part of an avalanche~~. We similarly assumed that the boundary of 'rough snow' should be equal to 1 for ~~both~~ the 'snow' and 'buffer' classes, ~~which means~~so that the segment lies completely within these two classes; ~~however, this time we combined it with~~we used a SD_{NDWI} of >=0.75 to include the segment as 'rough snow'. The thresholding value of the SD_{NDWI} ~~was taken~~came from the most pronounced changes in the ~~data~~ histogram ~~data~~by analysing the rapid change in the counts of values on the histogram. We ~~than~~ then merged the segments into the 'rough snow' class, ~~with the~~except ~~for~~in of areas <62.5 m² (area <1,000 pixels; Fig. 4), assuming that ~~they~~ these ~~were~~ are too small to represent an avalanche. We chose ~~The~~this minimum area threshold, ~~the same as in previous steps, we selected~~ based on visual checks and a ~~by performing~~ trial-and-error sensitivity analysis in the eCognition software.

4.1.3 Buffering

In further steps, we split all segments from the 'rough snow' class into smaller pieces, to reduce their ~~artificially~~ complex shapes with bigger and more compact parts connected to neighbours by only a few pixels. In most cases, only some of these complex shapes represented an avalanche, whereas the remainder was 'rougher snow' due to vegetation effects. To simplify

5   these shapes into separate parts, we used buffering to reclassify pixels from the 'rough snow' class as snow that had less than four neighbours classified as 'rough snow' (Fig. 4), repeating this step for both the 'snow' and 'rough snow' classes. To avoid undue growth of spurious pixels, we narrowed down the process to only pixels adjacent to at least one pixel classified as 'rough snow' (Fig. 4).

4.1.4 Neighbourhood analysis

10   At this stage, our classification still contained many misclassified parts of avalanches containing effects of vegetation, soil, or rocks. To assign these parts to the 'rough snow' class, we first reclassified all pixels from 'unclassified', 'buffer', 'vegetation', and 'dark objects' with brightness >3,000 and NDVI <140 into a new 'temp' class (Fig. 4). The brightness thresholds ~~for brightness we~~ ~~took by performing the~~ derived from a sensitivity analysis, in analogy to previous steps~~, which~~ ~~the results comparing to the accuracy in detecting each individual avalanche are shown in~~ (Fig. 6). After merging the

15   segments into the 'temp' class, we reclassified all segments bordering 'rough snow' <0.01 into 'unclassified' and discarded these from the analysis. For the remaining segments, after returning to chessboard segmentation to allow us to once again operate on single pixels, we reclassified all pixels from the 'temp' class sharing a border with 'rough snow' with at least two pixels (Fig. 4) iteratively. Because every pixel had only four neighbours, the assumption concerning two pixel neighbours stopped the infinite loop after only a few repetitions. To allow the inclusion of additional pixels, we therefore decreased the

20   threshold on the relative border to >=0.25 and performed the process twice more to allow more segments that were sharing the boundary with at least one 'rough snow' segment to be included in the 'rough snow' class. We did not run this process iteratively, because it would have reclassified all pixels assigned as 'temp', and our aim was to increase only compact 'rough snow' areas. Next, we once more applied an infinite loop regarding the relative border to 'rough snow' of >=0.5, to increase previously ~~detected~~ ~~snow~~ avalanches. These steps were crucial in closing areas inside the ~~snow~~ avalanches that due to the

25   values of brightness and NDVI, were not assigned in the previous steps to the 'rough snow' class.

4.1.5 Adding small gaps inside avalanches

Finally, we focused on filling gaps inside the detected snow avalanches. We reclassified and merged gaps into avalanches by verifying their geometrical relation to 'rough snow'. After checking the layer statistics for every segment, we built an assumption that if a segment was completely within the 'rough snow' and its area was <1,000 pixels (62.5 m²), it was to be

30   automatically reclassified into 'rough snow' (relative border to 'rough snow' = 1 and area <1,000 pixels; Fig. 4). Segments with >1,000 pixels reclassified into 'rough snow' were expected to fulfil additional rules concerning their roughness,

brightness, occurrence of snow, and vegetation, because they may have represented a convex-upward form that could stay intact during ~~the~~ avalanch~~eing occurrence~~ing. Only segments with a high snow roughness that did not represent~~ed~~ vegetation and were not too dark were added to the 'rough snow' class (relative border to 'rough snow' = 1, SD_{NDWI} >0.7, NDWI >127, NDVI <140, and brightness >2,500; Fig. 4). The threshold selected for brightness is lower than in ~~the~~ previous steps, because

5   some parts of snow avalanches were a little darker~~.;~~ ~~therefore~~hence ~~applying~~using the same threshold would ~~not allow their correct assignment to a proper class~~cause unwanted misclassifications. We exported all extracted snow avalanches into polygon shapefiles and compared these visually and quantitatively with manually-mapped reference data. The visual interpretation was important for verifying the distribution of errors and the completeness of classified avalanches.

4.2 Generating reference data

10   We created reference data by manually digitising avalanches from the images in ArcMap 10.3 software at scales between 1:800 and 1:1,500, depending on the complexity of the ~~mapped~~ avalanche debris. Manually digitising each area affected by the avalanche was necessary because avalanches occur only during the winter season when snow cover occurs, making them temporal events; the marks of their existence disappear when the snow cover melts, and therefore no complete reference data are available. Avalanches that were overlapping or bordering others were counted as one~~. This means that~~; some polygons

15   ~~representing an avalanche in fact covered~~contained several smaller avalanches that had occurred in succession; in most cases, it was impossible to assess their relative sequence (Fig. 3). We mapped a total of 2,200 avalanche polygons for the data acquired in the winter of 2012/13, ~~obtaining~~ covering 13.6 km² of avalanche terrain or 6% of the study area. The~~se~~ reference data~~, as well as automatically extracted avalanche polygons,~~ are available online (https://uni-potsdam.maps.arcgis.com/apps/Cascade/index.html?appid=3b5ac4491b59480c8c6016139f285e88)~~, and. We~~ we used the~~m~~

20   ~~reference data~~ to estimate several ~~classification accuracy~~performance metrics, including Type I, Type II, and total errors (Sithole and Vosselman, 2004), overall, user's, and producer's accuracies (Congalton, 1991), Cohen's kappa (Cohen, 1960), and F-Score. ~~Finally, w~~We arbitrarily selected three 4-km² training sites (Fig. 3) for our OBIA algorithm, and reported these eight performance metrics for a larger (226.3 km²~~)~~ test area.

5 Results

25   **5.1 Estimated accuracy**

Our algorithm classified 10.7 km² as avalanche debris, which is 78.7% of the total area of the reference data mapped for the winter of 2012/13. Overall, 1,648 out of 2,200 avalanches were correctly identified; 1,126 were detected in terms of more than half their area, and for 615 avalanches this detection rate was >80%. These classified avalanches were spread out evenly throughout the study area (Fig. 5). Visual checks of the classification indicated that the runout zones were detected most

30   reliably, whereas the release zones were the most problematic. Tracks were detected mostly correctly where small patches of vegetation or soil were near or in the avalanche tracks. The highest estimated precision in detecting avalanche boundaries

8

was in runout zones adjacent to smooth snow or ~~unvegetated~~ bare slopes. In some locations, a clear distinction between avalanche debris and smooth snow was ~~not possible~~difficult, especially for older deposits or snow drifts. Fresh avalanches were also detected with higher accuracy than ~~were~~ older and blurred ones.

The performance metrics estimated for our training sites were overall and user's accuracies of >0.9, and a Cohen's kappa of >0.8 (Tab. 1). The accuracy for the testing area yielded lower performance metrics, with a user's accuracy of 0.78. The producer's accuracy showed that 61% of the total avalanche area in the tested data was correctly identified. The overall Type II error was very low, indicating that few objects were falsely classified as avalanches, whereas the high Type I error showed that many, mostly old, snow avalanches remained undetected.

Table 1: Performance metrics estimated for 4-km² training sites 1 – 3 (see Figures 3 and 5) and for the entire study area covering 226.3 km².

	Type I error	Type II error	Total error	Overall accuracy	Producer's accuracy	User's accuracy	Cohen's kappa	F-score
Site 1	0.23	0.01	0.05	0.95	0.77	0.91	0.81	0.83
Site 2	0.23	0.02	0.08	0.92	0.77	0.92	0.79	0.84
Site 3	0.16	0.02	0.05	0.95	0.84	0.93	0.85	0.88
Total	0.39	0.01	0.03	0.97	0.61	**0.78**	**0.67**	0.69

5.2 Influence of variables used for classification accuracy

We checked how brightness, NDVI, NDWI, and SD_{NDWI} derivatives evaluated for each classified avalanche affected the producer's accuracy. For each avalanche, we computed the mean for each derivative map using the values of all the pixels inside the avalanche. We found that avalanches that were extracted with the highest accuracy were generally also brighter (Fig. 6); an increase in producer's accuracy occurred with an increase in the mean brightness of avalanche. A similar, but, weaker correlation held for SD_{NDWI}. Neither NDVI nor NDWI had much of an influence on the classification accuracy. In addition, we verified if an avalanche's shape (roundness) and size (area) affected the detection rate. Similarly to NDVI and NDWI, however, we did not find any dependence (Fig. 6).

5.3 Topographic factors favourable for snow avalanches

We further analysed the topographic settings of the mapped avalanches. Most avalanches (1,422 out of 2,200) occurred between 1,900 and 2,600 m a.s.l., with a mode of approximately 2,400 m a.s.l., on slopes that were 20 – 40° (Fig. 7). ~~The slope of occurred avalanches indicates that applying~~Our results show that avalanches were also detected on slopes steeper than 35°, thus highlighting the limits of detection methods using arbitrary slope thresholds ~~threshold, which has been used for example by~~ (Bühler et al., ~~(2009) and~~; Vickers et al., ~~(2016), would discard some of avalanches occurring on our~~

~~research area. It concludes, that constructing our OBIA approach excluding slope related assumption was a proper way in the~~ ~~classification.~~ One hundred thirty nine out of 193 of the highest-lying avalanches (>2,800 m a.s.l.) were small or very small ~~events~~, according to the nomenclature of Bühler et al. (2009) and occurred in the southern part of our research area. Although the largest avalanches occurred below 2,400 m a.s.l., most affected northeastern and the southwestern slopes, thus

5 mimicking the major aspects of the mountain ranges (Fig. 3).

5.4 Density of avalanches

Using the reference data for the winter of 2012/13, we further analysed where most of the avalanches occurred. We computed centroid locations for each avalanche polygon, and estimated their spatial density using the ~~k~~*Kernel density* function in ArcGIS 10.3, with both point- and area-weighted inputs in a 2-km radius. We produced two maps because, due to

10 the occurrence of multiple avalanches, our input centroids did not represent the total number of events. Consequently, a point-weighted map could have underestimated the real avalanche density, whereas an area-weighted map avoids this issue. The selected 2-km size of the bandwidth was large enough to avoid reproducing the pattern of input avalanche centroids, and small enough to reduce the smoothing of the point information. We found that avalanches clustered largely in the south-eastern part of the study area (125 smaller avalanches in inset ~~1,~~on Fig. 8.a, of which 104 were <2,000 m²). The area-

15 weighted spatial density of avalanches was highest on the slopes of Fluela Schwarzhorn, Sentischhorn, and Wuosthorn. The biggest avalanche was ~~in inset 2 on Figure 8.b~~ had an area of ~390,000 m² (Fig. 8.b).

5.5 Automatic classification of snow-avalanche zones

Automatic delineation of release and runout zones from remote sensing data can be conducted by ~~verifying~~ using the elevation ~~values~~ range in each avalanche polygon ~~that represents an avalanche~~. The simplest approach uses ~~a~~ flow length,

20 which ~~represents~~ is the distance along the flow path inside the avalanche, as an indicator in evaluating the probability of the release and the runout zones. ~~Herein, w~~We suggest using the elevation values in an approach that allows exposure of local terrain height differences in the probability map. For example, applying the elevation for a release zone that occurs on very steep slopes determines that the highest probability will be represented only for a very small area, whereas on gentle slopes, the same probability will occur over a larger area, because the elevation differences are smaller than those of steep slopes.

25 We used a 2-m resolution digital surface model (DSM) derived from stereomatching of aerial ADS images to automatically detect release zones, tracks, and runout zones. Because we generated our reference data in the same resolution as the aerial images, we first resampled the DSM to 0.25-m resolution (Fig. 9). We then used the reference data to clip the DSM to each avalanche polygon using a key ID~~.~~; the DSMs acquired in this way were used to compute the probability of representing individual parts of an avalanche. We assumed that the maximum (minimum) elevation on which the snow avalanche occurs

30 has the highest probability of being a release (runout) zone. We estimated the probability by normalising elevation data (Fig. 9) by subtracting the minimum value of DSM inside this avalanche from DSM pixel value, and dividing by the difference between the maximum and the minimum values inside each avalanche. Such stretching returns a 0 – 1 probability map for

10

release areas. The thresholds in classifying the track are relative; we applied 0.3 – 0.8 as the thresholds, because most of our 2,200 manually-digitised polygons represented loose snow avalanches where the release zone constituted a small part, and the runout zone compared to the track constituted an even smaller part. When dealing only with loose avalanches, the upper threshold for release areas may be set to a probability equal to 0.9. The selected thresholds worked well for single

5 avalanches, where both the release zone and the runout zone were assigned correctly (Fig. 9). For avalanches with more complex shape and multiple avalanche arms (e.g. avalanche 1,207 on test site 2; Fig. 9), however, the release zones remain undetected, with the exception of only the highest-lying release areas.

6 Discussion

Automatic mapping of avalanches is crucial in mountainous regions to delineate susceptible areas and to produce cadastres

10 for validating avalanche forecasting and hazard maps. Previous studies have demonstrated that a combination of aerial images and digital elevation models (Bühler et al., 2009), or aerial (Lato et al., 2012) and satellite (Larsen et al., 2013), or SAR images (Vickers et al., 2016) ~~allow~~ favours the automatic ~~the~~ detection~~on~~ of ~~of~~ snow ~~-~~ avalanche deposits; however, identifying release zones and tracks remains challenging. We ~~have~~ proposed~~d~~ an OBIA algorithm that tracks release zones, tracks, and runout zones in NIR images. We recommend using normalised derivatives NDWI and NDVI, instead of

15 brightness, for classifying water (snow) and vegetation, respectively, because the thresholds allowing for classification of water and vegetation in these indices are stable (around zero) and widely applicable. Our approach ~~upon~~ expands upon previous work, as we have considered potential ~~snow~~ avalanches in snow-covered areas only (NDWI >127), whereas others have used brightness thresholds (Lato et al., 2012), with the ~~danger~~ pitfall of including other objects with similar values to those of snow. Rivers and lakes have similar values of NDWI, but are less bright than snow, we therefore combined NDWI

20 and brightness in our model. We suggest using SD_{NDWI} to trace rough snow or avalanche debris. Estimated accuracy was high for detecting avalanches using the test data, and the random spread of avalanches detected with high accuracy through the whole research area suggests that the assumptions in our approach are broad enough to be applied for a large area.

To increase the accuracy of true positives in detecting release zones in our classification, a customised threshold for every data tile can be applied. In any case, we focused on developing a method that is transferable and works well for a greater

25 area. We also wanted to test the usability of the image derivatives and verify how much information we could obtain from them when detecting snow avalanches. We therefore did not change any parameter or its threshold when applying the algorithm for other data tiles (Lato et al. 2012). We implemented several steps that allowed the classification of different avalanche scenarios that occur in diverse topographic conditions, such as single vs. multiple avalanche; small vs. large avalanche; avalanche revealing the ground or vegetation vs. avalanche that does not reveal the ground or vegetation;

30 avalanche that is blocked by vegetation or other objects having high roughness vs. avalanche that is not adjacent to any rough object. These topographic conditions influence the appearance of avalanches on images; therefore, the number of steps in our OBIA workflow is large. We found that OBIA is useful for complex shapes, because it allows the implementation of assumptions regarding each different situation. Additionally, this OBIA algorithm may be used and modified according to

11

the need of the user and the data. For example, some steps in our OBIA workflow may be omitted, such as those used for filling the gaps inside the avalanche. If the data do not contain avalanches that reveal bare ground or vegetation, these steps may not improve the classification, because there will not be any gaps to reclassify. Additionally the number of loops in the shrinkage and growth steps depend on pixel size. With our 0.25-m resolution NIR images, we shrank and increased the

5   segments of 1.25-m, which was sufficient to split segments into small parts; with higher (lower) resolution, the number of loops should be increased (decreased) accordingly. Previous studies have reported very high accuracy in separating vegetation and water using vegetation and water indices (Townshend and Justice, 1986; Ji et al., 2009); therefore, we assume that the NDVI, and NDWI thresholds are stable and may be easily transferable to other areas and data. Despite this, the thresholds for SD_{NDWI} should always be verified by analysing the mode of data distribution on a histogram of the SD_{NDWI},

10   testing the thresholding of data samples, and checking the thresholding results visually. We are aware that such visual checking may introduce some bias, although so far it is the most common way in OBIA to find the most suitable threshold.

The size and the shape of avalanches did not influence the classification accuracy if they were bigger than 2,000 pixels, which is what we regard as the minimum detectable size in our OBIA algorithm. We selected the minimum number of pixels necessary by analysing the reference data, where avalanches smaller than 2,000 pixels constituted only 7.5% of the total

15   number of avalanches, and by taking into account the classification of avalanches with respect to their size. According to Bühler et al. (2009), very small avalanches are <100 m², which coincides with 2,000 pixels (125 m²). The parameters playing the biggest roles were brightness and the SD_{NDWI} (Fig. 6). Visual inspection showed that the easiest avalanche part to detect was the runout zone and the most difficult was the release zone, because release zones were not usually rough enough or did contain outcrops of the vegetation or bare ground. The correct detection of the track depended mostly on snow roughness

20   and depth. In a case of low roughness values or a very thin snow cover revealing the ground and the vegetation, the track was not detected correctly or not detected at all.

Our method ~~allowed to~~ correctly identif~~ied~~y 61% of the total avalanche area in ~~research~~the study area. ~~This value comes mostly from the fact that the~~, mainly because of heterogeneous avalanche debris ~~properties of some part of avalanches differ from the majority of avalanches pixels~~ with dark or smooth patches, and because of differing deposit ages, and hence

25   differing surface roughness~~, for example they are too dark, or too smooth~~. ~~In addition, many of avalanches were nor fresh and their properties differ significantly from the properties of fresh avalanches what makes their simultaneous classification more difficult.~~ The most difficult to classify were old avalanches where the bare ground cropped out or where vegetation occurred in the path of the avalanche. These avalanches did not meet the assumptions in our OBIA protocol and could not be classified correctly because they were not rough enough, or were too vegetated or dark due to thin snow cover. Here a

30   solution may be ~~applying~~using different thresholds for old ~~vs.~~and fresh avalanches to ~~fit~~optimise our OBIA approach ~~to such different types of objects~~. We tested different thresholds for the input layers and different neighbourhood assumptions ~~in order~~ to include these avalanches; however, this resulted in more false positives, so the cost of correct classification of these avalanches was higher than the benefit. We therefore decided to stay with the same workflow and thresholding shown in Figure 4 for the whole study area.

12

Errors in automatic classification also occurred due to data tiling. Using an Intel Xeon E5-2667U processor with 256GB RAM memory, we were able to run our OBIA algorithm only for tiles of 6.25 km² (10,000 columns x 10,000 rows), requiring a computing time of ~30 minutes for each tile; our test in executing bigger tiles ended in crashing the computation in the eCognition software. In many cases the avalanches were therefore split into two or more neighbouring tiles, which

5   influenced the correct detection of avalanche parts, especially those where the avalanches were spread across tiles or had small gaps.

Avalanches often have a tongue shape that may also be used as a property for classifying. In our test area, however, using such information may be insufficient, because we are dealing with avalanches repeated in the same location and merged avalanches, which have more complex shapes. In such cases, a new avalanche may partly cover previous avalanche, making

10   it difficult to distinguish them. In addition, a few avalanches may simply have a more complex and difficult-to-interpret shape. Detailed enough ~~Another advantage~~maps of snow thickness may help to improve the classification accuracy ~~may be the use of snow thickness maps,~~ ~~which could help~~and perhaps enable separating ~~to distinguish between~~ release zones, tracks, and runout zones from the surrounding snow cover. ~~In our OBIA approach we did not applied the thickness map, due to their low quality; with better quality data it should be considered in future research, however.~~

15   Our estimated avalanche density map and information concerning the most common prevalence of avalanches may be used to help generate a hazard map in mountainous areas. The most crucial issue, however, is whether the size or the number of avalanches is more important in such mapping. As mentioned by Eckerstorfer et al. (2016), only very small avalanches cause less damage; therefore, we suggest weighting the density map according to avalanche size to give more information regarding the degree of danger in a specific area.

20   Continuation of our research should contain verification of the transferability of our OBIA algorithm to data from different winter years for the area of Davos, Switzerland, which we have in our repository. ~~In addition, f~~Future tests involving diverse areas across the World and different types of data (e.g. UAV, radar, and seismic) are desirable. An approach for distinguishing single and repeat avalanches at the same area should be developed, because it would give more detail about the quantity and the frequency of avalanches in a given area.

25   **7 Conclusions**

We have presented an automatic object-based image analysis (OBIA) approach to detecting snow avalanches and their release, track, and deposition zones for a large region from ADS80 NIR aerial images. We used image-derived parameters, including the normalised difference vegetation index (NDVI), and the normalised difference water index (NDWI) and its standard deviation (SD_{NDWI}), to separate vegetation, snow, and rough snow representing avalanche debris, respectively. We

30   applied buffering assumptions relying on the local neighbourhood of segments to remove salt-and-pepper noise and objects that were falsely assigned as avalanches, using the thresholding of derivatives obtained from NIR images. For an area of 226.3 km², our algorithm achieved producer's and user's accuracies of 0.61 and 0.78, respectively, and a Cohen's kappa of 0.67. Our algorithm uses only information taken from these images, from which the evaluated NDVI and NDWI indices are

normalised, helping to transfer their thresholds to other areas. Our approach contains only three fixed parameters (NDVI, NDWI, and SD$_{NDWI}$) and two changeable parameters (brightness and segment area). The first depend on the spectral characteristics of images and the latter one depend on the resolution of images. Our OBIA workflow is not sequential when compared to that of Lato et al. (2012); the segments that do not fulfil assumptions in one step can still be considered as potential snow avalanches in the next steps. To assign potential snow avalanches, our method takes into account only the rough snow, which allows more reliable detection of avalanches. Our probability approach determines, in an automatic way, the highest and the lowest parts of the avalanche, and thus its release zone, track, and runout zone, which allows easy analysis of the topographic condition of areas where the avalanche starts and where the snow is deposited. For multiple avalanches with a complex shape, our probability map may not be sufficient to correctly identify all release and runout zones, but for single avalanches, it gives valuable results. For some avalanches, we were not able to judge visually if they were single or multiple; a discussion of this topic should therefore be undertaken. The probability approach may be used for any other mass movement landforms, such as landslides, to delineate their release and deposition zones. In the future, we plan to validate our snow avalanche algorithm for ADS data, which we have for other winters, and to verify its transferability to other NIR images, because successful results in this matter may offer a chance to improve hazard maps and avalanche forecasting in Switzerland.

References

Badoux, A., Andres, A., Techel, F. and Hegg, C.: Natural hazard fatalities in Switzerland from 1946 to 2015. Nat. Hazards Earth Syst. Sci., 16, 2747–2768, doi:10.5194/nhess-2016-232, 2016.

Bagli, S. and Schweizer, J.: Characteristics of wet-snow avalanche activity: 20 years of observations from a high alpine valley (Dischma, Switzerland). Nat. Hazards, 50, 97–108, doi:10.1007/s11069-008-9322-7, 2009.

Bründl, M., Etter, H. J., Steiniger, M., Klingler, C., Rhyner, J. and Ammann, W. J.: System Sciences IFKIS – A basis for managing avalanche risk in settlements and on roads in Switzerland. Nat. Hazards Earth Syst. Sci., 4, 257–262, doi:1684-9981/nhess/2004-4-257, 2004.

Bühler, Y., Adams, M. S., Bösch, R. and Stoffel, A.: Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): Potential and limitations. Cryosphere, 10, 1075–1088, doi:10.5194/tc-10-1075-2016, 2016.

Bühler, Y., Adams, M. S., Stoffel, A. and Boesch, R.: Photogrammetric reconstruction of homogenous snow surfaces in alpine terrain applying near-infrared UAS imagery surfaces in alpine terrain applying near-infrared UAS imagery. Int. J. Remote Sens., 0, 1–24, doi:10.1080/01431161.2016.1275060, 2017.

Bühler, Y., Hüni, A., Christen, M., Meister, R. and Kellerberger, T.: Automated detection and mapping of avalanche deposits using airborne optical remote sensing data. Cold Reg. Sci. Technol., 57, 99–106, doi:10.1016/j.coldregions.2009.02.007, 2009.

Bühler, Y., Meier, L. and Ginzler, C.: Potential of operational high spatial resolution near-infrared remote sensing instruments for snow surface type mapping. IEEE Geosci. Remote Sens. Lett., 12, 821–825, doi:10.1109/LGRS.2014.2363237, 2015.

Castebrunet, H., Eckert, N. and Giraud, G.: Snow and weather climatic control on snow avalanche occurrence fluctuations over 50 yr in the French Alps. Clim. Past, 8, 855–875, doi:10.5194/cp-8-855-2012, 2012.

Christen, M., Kowalski, J., and Bartelt, P.: RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. Cold Reg. Sci. Technol., 63, 1–14, doi:10.1016/j.coldregions.2010.04.005, 2010.

Cohen, J.: A coefficient of agreement for nominal scales. Educ. Psychol. Meas., 20, 37–46, 1960.

Congalton, R. G.: A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ., 37, 35–46, doi:10.1016/0034-4257(91)90048-B, 1991.

Eckerstorfer, M. Bühler, Y., Frauenfelder, R. and Malnes, E.: Remote sensing of snow avalanches: Recent advances, potential, and limitations. Cold Reg. Sci. Technol., 121, 126–140, doi:10.1016/j.coldregions.2015.11.001, 2016.

Eckerstorfer, M. and Malnes, E.: Manual detection of snow avalanche debris using high-resolution Radarsat-2 SAR images Radarsat-2 SAR images. Cold Reg. Sci. Technol., 120, 205–218, doi:10.1016/j.coldregions.2015.08.016, 2015.

Ji, L., Zhang, L. and Wylie, B.: Analysis of dynamic thresholds for the normalized difference water index. Photogramm. Eng. Remote Sens., 75, 1307–1317, doi:10.14358/PERS.75.11.1307, 2009.

Larsen, S.Ø., Salberg, A. and Solberg, R.: Automatic avalanche mapping using texture classification of optical satellite imagery. In: Lasaponara, R., Masini, N., Biscione, M. (Eds.), EARSeL, 2013. Towards Horizon 2020, pp. 399–410, 2013.

Lato, M. J., Frauenfelder, R. and Bühler, Y.: Automated detection of snow avalanche deposits: Segmentation and classification of optical remote sensing imagery. Nat. Hazards Earth Syst. Sci., 12, 2893–2906, doi:10.5194/nhess-12-2893-2012, 2012.

McClung, D. and Schaerer, P.: The Avalanche Handbook. The Mountaineers Book, Seattle, 2006.

McClung, D.M.: Avalanche character and fatalities in the high mountains of Asia. Ann. Glaciol. 57, 114–118, doi:10.3189/2016AoG71A075, 2016.

McFeeters, S. K.: The use of the normalized difference water index (NDWI) in the delineation of open water features. Int. J. Remote Sens., 17, 1425–1432, doi:10.1080/01431169608948714, 1996.

15

Schweizer, J. and Jamieson, J. B.: Snow cover properties for skier triggering of avalanches. Cold Reg. Sci. Technol., 33, 207–221, doi:10.1016/S0165-232X(01)00039-8, 2001.

Schweizer, J., Jamieson, J. B. and Schneebeli, M.: Snow avalanche formation. Rev. Geophys., 41, 1–25, doi:10.1029/2002RG000123, 2003.

5 Sithole, G. and Vosselman, G.: Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS J. Photogramm., 59, 85–101, doi:10.1016/j.isprsjprs.2004.05.004, 2004.

Techel, F., Jarry, F., Kronthaler, G., Mitterer, S., Nairz, P., Pavsek, M., Valt, M. and Darms, G.: Avalanche fatalities in the European Alps: Long-term trends and statistics. Geogr. Helv., 71, 147–159, doi:10.5194/gh-71-147-2016, 2016.

Techel, F., Zweifel, B. and Winkler, K.: Analysis of avalanche risk factors in backcountry terrain based on usage frequency
10 and accident data in Switzerland. Nat. Haz. Earth Syst. Sci. 15, 1985–1997, doi:10.5194/nhess-15-1985-2015, 2015.

Townshend, J. R. G. and Justice, C. O.: Analysis of the dynamics of African vegetation using the normalized difference vegetation index. Int. J. Remote Sens., 7, 1435–1445, doi:10.1080/01431168608948946, 1986.

Varma, M. and Zisserman, A.: A statistical approach to texture classification from single images. Int. J. Comput. Vision, 62, 61–81, doi:10.1007/s11263-005-4635-4, 2004.

15 Vickers, H., Eckerstorfer, M., Malnes, E., Larsen, Y. and Hindberg, H.: A method for automated snow avalanche debris detection through use of Synthetic Aperture Radar (SAR) imaging. Earth Space Sci., 3, 446–462, doi:10.1002/2016EA000168, 2016.

Wilhelm, C., Wiesinger, T., Bründl, M. and Ammann, W.: The avalanche winter 1999 in Switzerland - An overview. Proc. of the 2000 International Snow Science Workshop. Montana, USA, pp. 487–494, 1999.
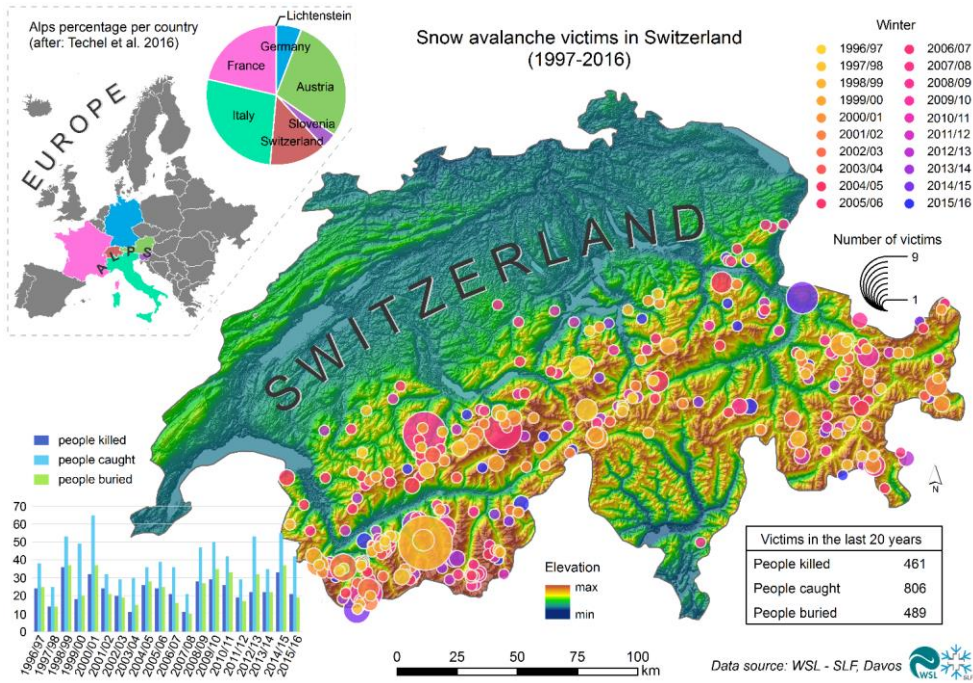
20

Figure 1: Snow avalanche accidents with victims in Switzerland in the winters of 1996/97 to 2015/16, and the Alps percentage per country. Data from: Swiss Federal Institute for Snow and Avalanche Research (WSL-SLF), Davos, Switzerland.
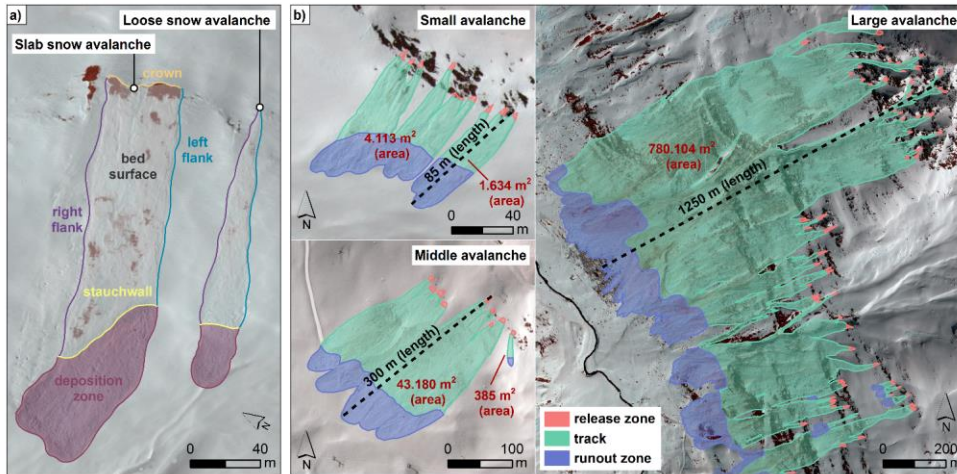
Figure 2: a) Types of avalanche release: slab avalanches and loose avalanches, with marked avalanche body parts (crown, bed surface, stauchwall, deposition zone, and right and left flanks; Schweizer et al. 2003); b) Avalanche classification with respect to length: small, middle, and large [m], with the avalanche area [m²] and marked parts of avalanches: release zone, track, and runout zone. Data from: WSL Institute for Snow and Avalanche Research (SLF), Davos, Switzerland.
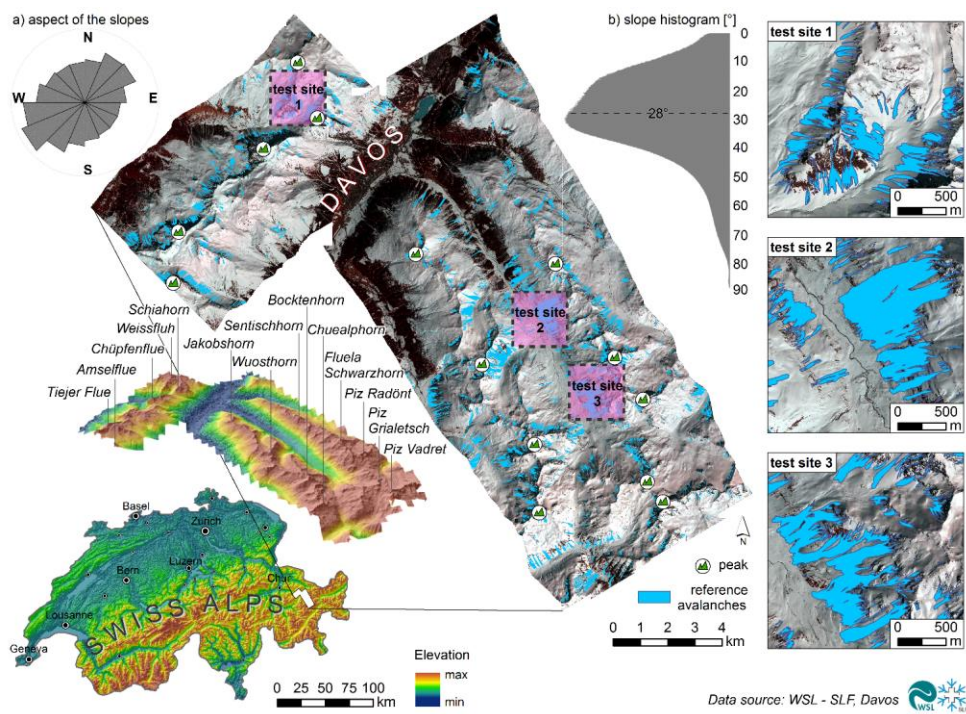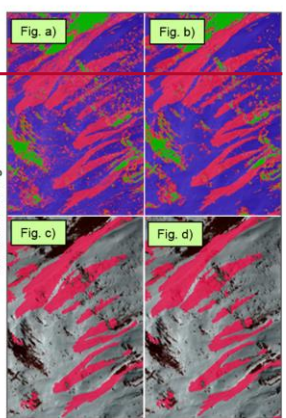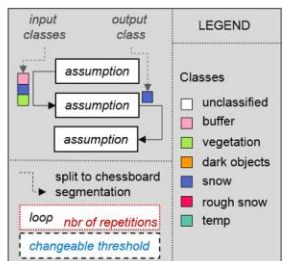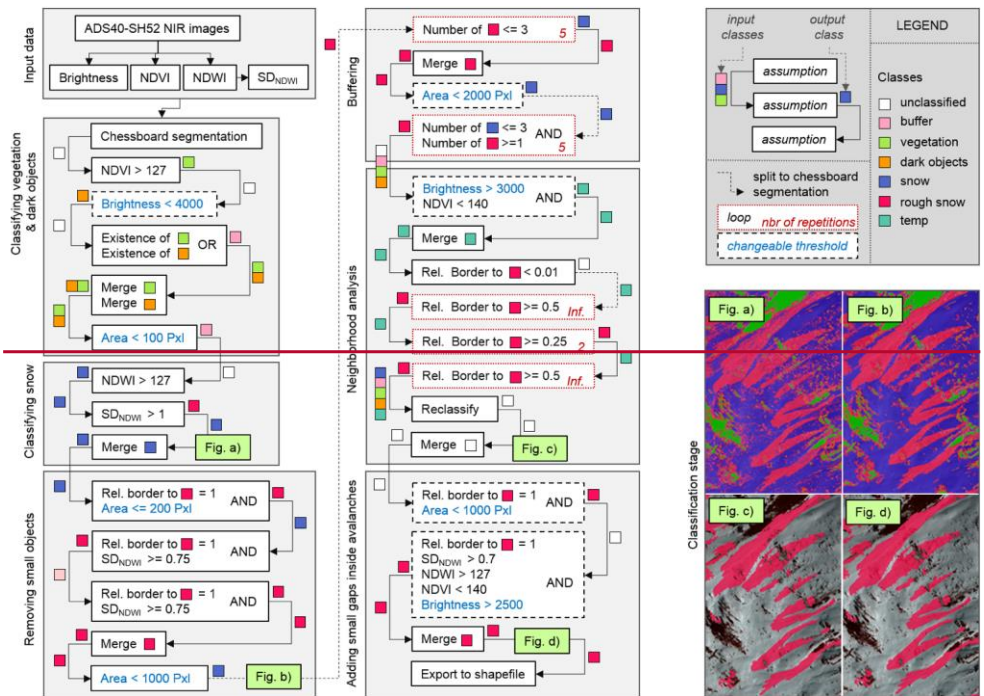
Figure 3: Topographic setting of research area and ADS 80 NIR aerial images with test sites (1 – 3) and digitally mapped snow avalanches; different colours mark different snow avalanche polygons. Inset histograms show a) distribution of local slope and b) the main slope aspect.

Input data

ADS40-SH52 NIR images

Brightness | NDVI | NDWI | SD_{NDWI}

Classifying vegetation & dark objects

Chessboard segmentation

NDVI > 127

Brightness < 4000

Existence of / Existence of — OR

Merge / Merge

Area < 100 Pxl

Classifying snow

NDWI > 127

SD_{NDWI} > 1

Merge — Fig. a)

Removing small objects

Rel. border to ☐ = 1 AND Area <= 200 Pxl

Rel. border to ☐ = 1 AND SD_{NDWI} >= 0.75

Rel. border to ☐ = 1 AND SD_{NDWI} >= 0.75

Merge

Area < 1000 Pxl — Fig. b)

Buffering

Number of ☐ <= 3   5

Merge

Area < 2000 Pxl

Number of ☐ <= 3 AND Number of ☐ >=1   5

Neighbourhood analysis

Brightness > 3000 NDVI < 140   AND

Merge

Rel. Border to ☐ < 0.01

Rel. Border to ☐ >= 0.5   Inf.

Rel. Border to ☐ >= 0.25   2

Rel. Border to ☐ >= 0.5   Inf.

Reclassify

Merge — Fig. c)

Adding small gaps inside avalanches

Rel. border to ☐ = 1 AND Area < 1000 Pxl

Rel. border to ☐ = 1 SD_{NDWI} > 0.7 NDWI > 127 NDVI < 140 Brightness > 2500   AND

Merge — Fig. d)

Export to shapefile

LEGEND

input classes   output class

assumption
assumption
assumption

→ split to chessboard segmentation

loop   nbr of repetitions

changeable threshold

Classes

☐ unclassified
☐ buffer
☐ vegetation
☐ dark objects
☐ snow
☐ rough snow
☐ temp

Classification stage
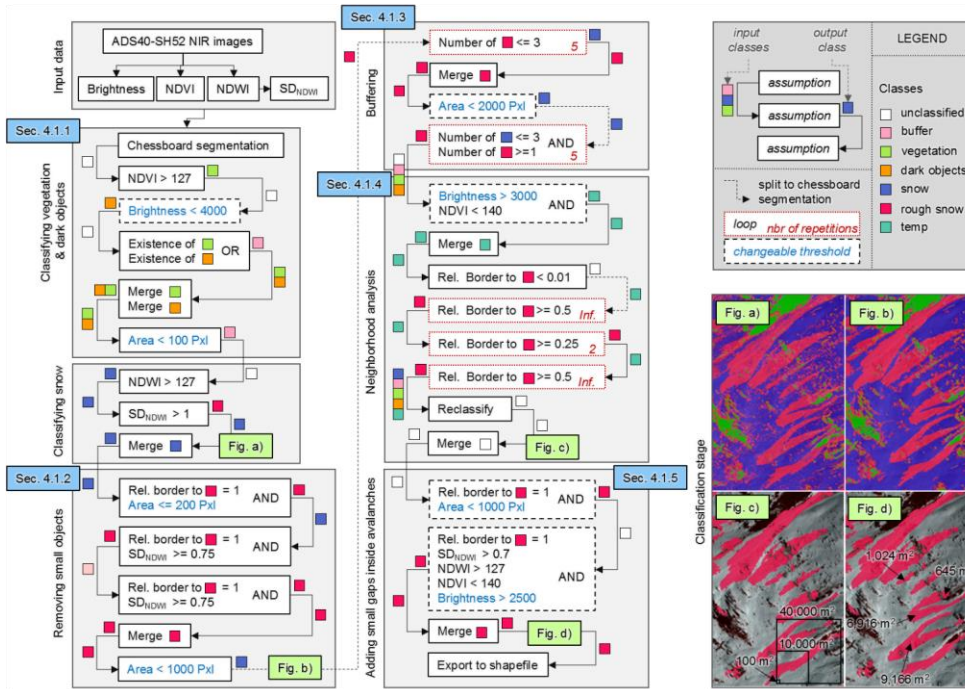
Fig. a)   Fig. b)

Fig. c)   Fig. d)

20

Figure 4: Workflow for classifying snow avalanches with object-based image analysis (OBIA). The white boxes indicate the classification; dashed outlines feature the number of reiterations (red), and the local decision boundaries show change (blue). The colours of squares are coded to the input and output class in each step. Figs. a), b), c), and d) represent visual results of the sub-step classification. The squares and numbers on Figures c and d represent a sample segment size. Sec. 4.1.1 – 4.1.2 of the figure are the numbers of subchapters in the article.
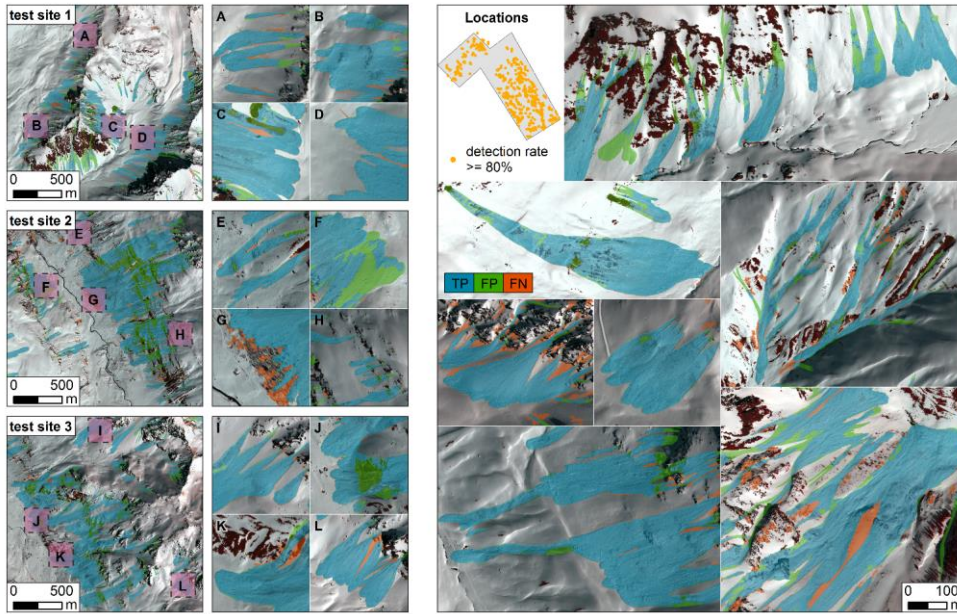
Figure 5: Accuracy assessment of OBIA classification for training (test sites 1 – 3) and testing data together; avalanche debris are shaded with orange colour, where they were detected with an accuracy >=80%. TP = true positive; FP = false positive; FN = false negative.
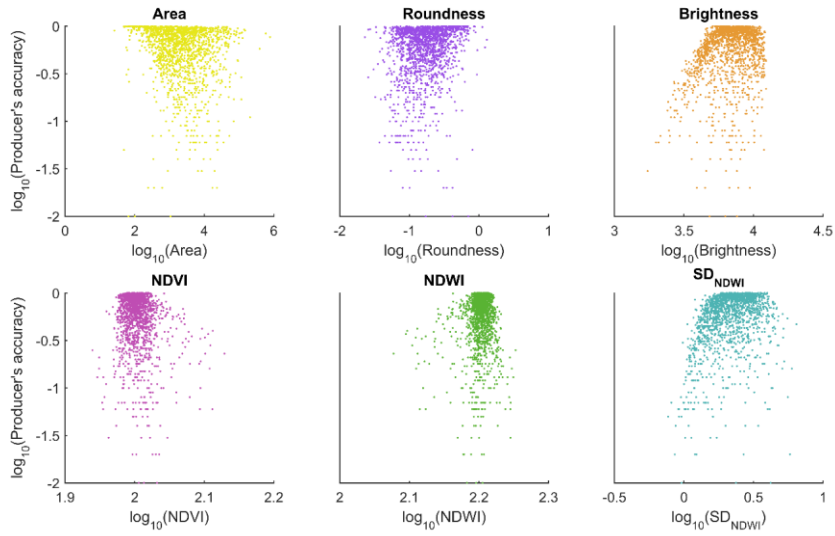
Figure 6: The role of avalanche area, roundness, brightness, normalised difference vegetation index (NDVI), normalised difference water index (NDWI), and standard deviation of normalised difference water index (SD$_{NDWI}$) in the estimated accuracy, when detecting snow avalanches with our OBIA approach.
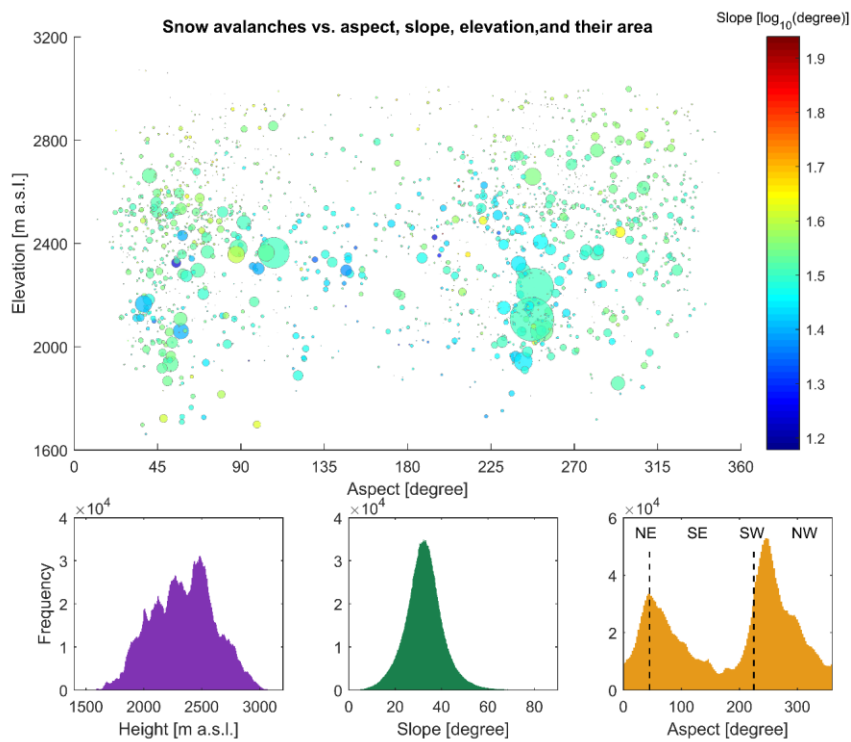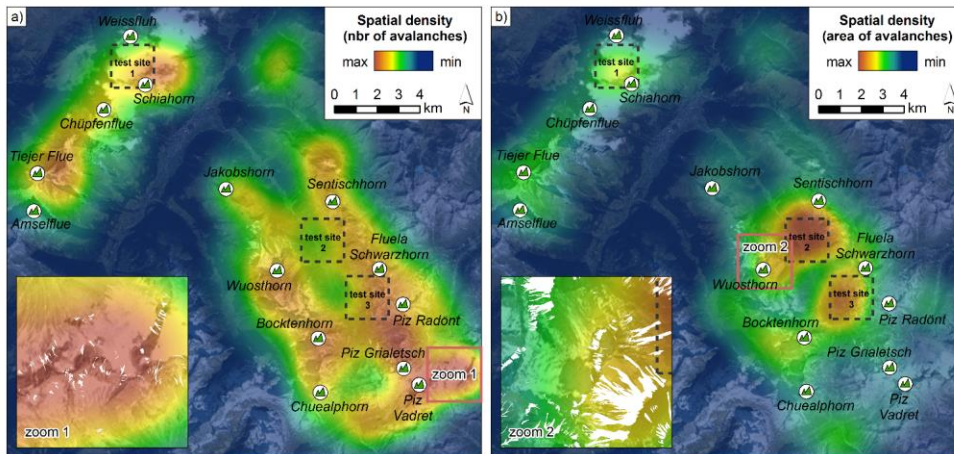
Figure 7: Elevation, slope, and aspect of mapped reference avalanches with marked northeast and southwest directions in the winter of 2012/13. The values presented for every single avalanche represent the mean value of each pixel contained inside the reference avalanche polygon. Bubble size on the scatter plot is scaled to avalanche area.

Figure 8: Density maps showing the clustering of avalanches in the winter of 2012/13, with respect to their a) quantity, and b) size. The insets (zoom 1 and zoom 2) show the density maps with respect to manually-classified reference snow avalanche polygons shown in white. Outlined test sites 1 – 3 are the test sites that were used for developing our OBIA algorithm.
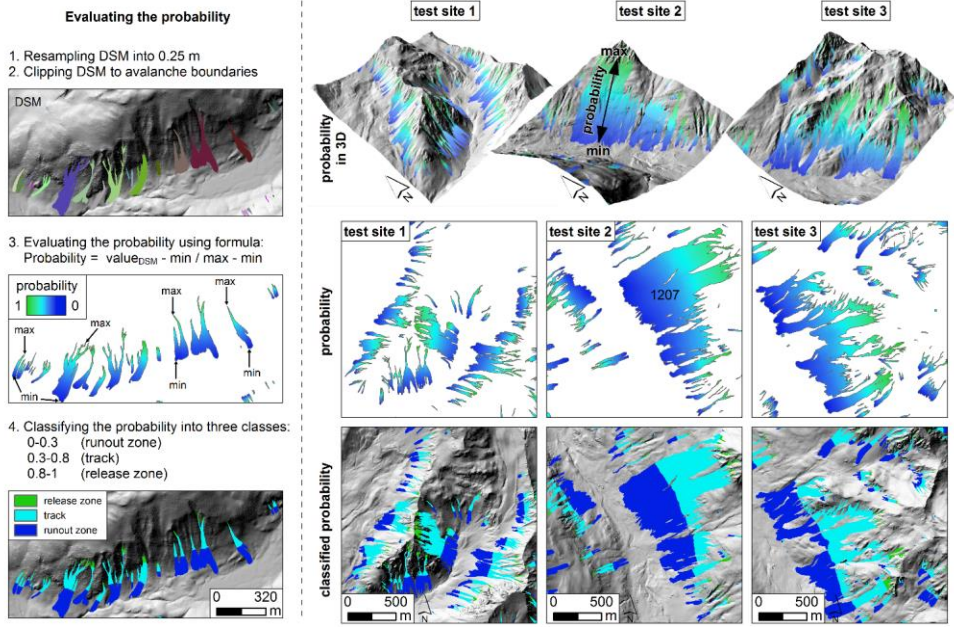
Figure 9: Estimating the probability of an avalanche release area. A probability close to one indicates a pixel representing a release zone, whereas a probability close to zero indicates a runout zone in automatically classified data.