Referee #1:

The authors thank Reviewer #1 for his comments on the manuscript. Please find our answers to the comments in green text below.

Thank you for the opportunity to the review this manuscript.

The authors develop an event-based inventory of shallow landslides for an area that includes forested and unforested zones. They find that some landslide characteristics (e.g., depth) host significant relationships with the presence/absence of forest and distance to individual trees. The authors conclude that these findings improve our understanding of how forests affect shallow landslide processes.

As written, it is sometimes difficult to discern if the main focus of this study is meant to be semi-automated landslide mapping methods or developing a better understanding of landsliding in context of vegetation. Substantial text is dedicated to describing the setup, tweaking, and performance of the segmentation algorithm. Yet, given the vegetation-oriented nature of the Abstract and Conclusion, I found that (1) relatively little text is developed to review the existing literature to shape the study design as part of the Introduction, (2) the analyses do not attempt to control for the effects of other landslide-relevant variables (e.g., topographic slope angle) inside/outside the forest cover mask as part of the Methods, and (3) process-based interpretations of the observed relationships between vegetation and landslides are underdeveloped as part of the Discussion.

Regarding Point #2, my impression is that authors have developed an internally consistent methodology that facilitates reproducible observations of differences in landslide dimension inside/outside the forest mask, but it is unclear why is there no attempt to “normalize” these comparisons in the context other spatially variable factors. The noted presence of steeper slopes in the forested versus unforested area (33 degrees versus 23 degrees, respectively; LN 104-105), for example, suggests that some consideration of landslide occurrence in the context of topographic slope angle, for areas with forest/no forest, should be considered. Or, it may suggest that the forest/no forest mask is not well suited for evaluating landslide characteristics in the context of vegetation for the study area. In the absence of an effort to deconvolve these kinds of competing factors, which the authors also note in the text (LN 419-420), it is difficult to evaluate if the conclusion (i.e., that the study provides “a better understanding of the roles of forest and how they affect the processes behind shallow landslides”) is supported by the results.

Sincerely,

Matthew A. Thomas

Thank you for raising these valid points and suggesting these interesting additional analyses.

Regarding point 1: According to the authors best knowledge all relevant papers focusing on the characteristics of landslides in relation to their forest cover, at least at a catchment scale level, were considered in the introduction. We did find one additional paper addressing the effect of forest cover on landslide characteristics (Schmidt et al., 2001), which will be added. To address the raised concern about the thematic focus of the paper, we will also add additional papers to the introduction and discussion that focus more generally on the interaction of vegetation related processes and landslides (mostly focusing on the occurrence of landslides) (such as Moos et al. (2016); Preti et al. (2013); Schwarz et al. (2010) and Sidle and Ochai (2006)).

Regarding point 2: Two additional analyses were performed to deconvolute the additional factors that might influence the differences in landslide characteristics.

(1) To investigate the effects of land-cover changes, we inspected historic orthophotos (starting from 1988 until 2015) to investigate if the locations of the landslides experienced significant land-cover changes in the 20 years leading up to the event in 2015 (such as deforestation or afforestation). The analysis found that of the 107 landslides located outside the 2013 forest mask, 47 landslide locations were previously located in the forest and clear-cut. Of the 48 landslides located inside the 2013 forest mask, 9 were located in previous grassland locations.

When these landslides are excluded from the analysis we still find similar results. The morphometric analysis now also shows statistically significant differences in the thickness of the landslides after forest mask classification with a p-value of 0.008 for the Mann Withney U test and a p-value of 0.007 for the t-test (Fig. 1).

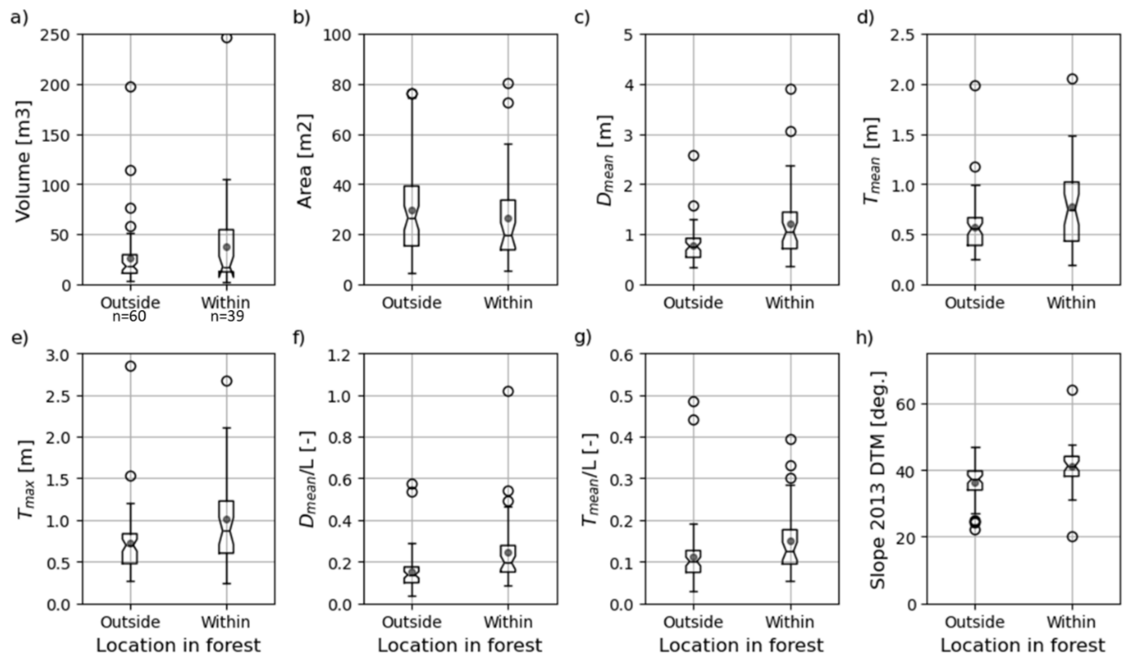

Figure 1: Updated analysis of the morphometric landslide characteristics, where the landslides at land-cover change locations are excluded.

The morphometric analysis with the forest structure parameters, as performed in the manuscript, also still shows a significant difference in the thickness values of the landslides using a classification by the average tree distance, with a p-value of 0.032. Since we already discussed internally as authors that the average minimum tree distance is a complex measurement that does not clearly represent the effect of a tree in close proximity to the scar, we also tested a classification by the minimum tree distance of the entire scar. Using this minimum tree distance metric, the distinction in scar thickness values between the classes becomes clearer, with a p-value of 0.006. A significant difference with a classification using the number of trees within a 10 m radius (tree density), could not be found. We will update the results section of the paper with these new results using the minimum tree distance.

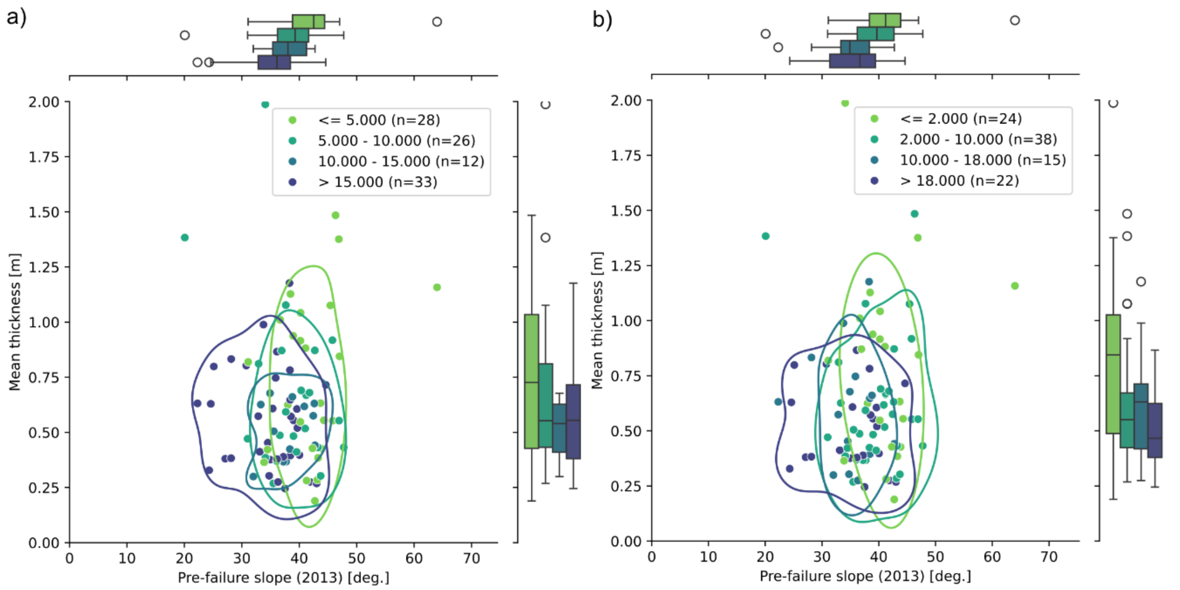

Figure 2: Updated analysis of relationships between "Mean thickness"(y-axis) – "pre-failure slope"(x-axis), colored and classified by "Mean tree distance" of the scars (≤5, 5-10, 10-15 and >15 m) (a) and colored by "Minimum tree distance" of the scars (≤2, 2-10, 10-18 and >18 m) (b).

The analysis of the landslide profiles shows similar results. Generally, the profiles are very similar to in the original manuscript, but the distinction between the profiles inside and outside the forest has become clearer.

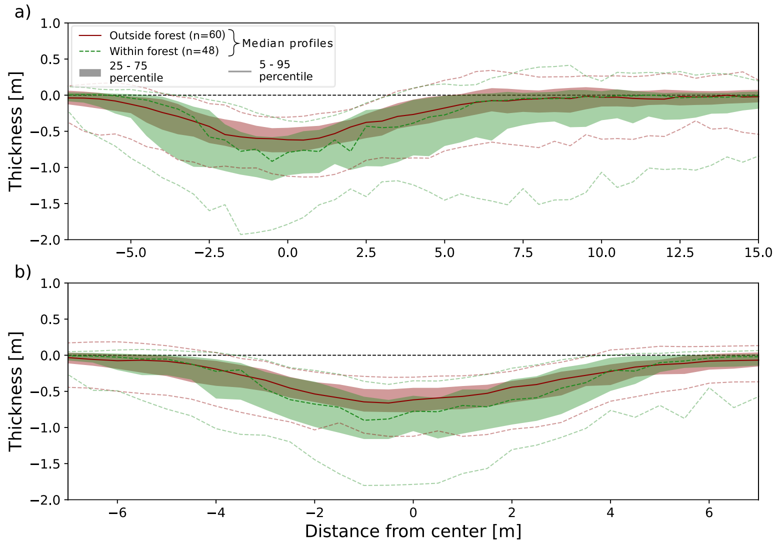

Figure 2: Topographic profiles from the thickness dataset of the scars within the forest (green) and outside the forest (red) along the aspect direction (a) and across the aspect direction (b). The x-axis indicates how far the point of the profile is located from the centroid of the scar.

(2) To deconvolute the effects of the slope from the found differences in landslide thickness, it was also investigated if a significant difference could be found with a classification by the pre-failure slope value. For this we also filtered out the landslides at locations of significant land-cover change. The ANOVA results show that a classification by slope does not result in significant differences between the thickness values in the slope classes (p-value of 0.69). This shows that it cannot be the effects of slope on the thickness values which are showing up in differentiation between the forest classes. The authors would like to stress that this does not rule out the effects of the slope on the occurrence of landslides, which is a different topic.

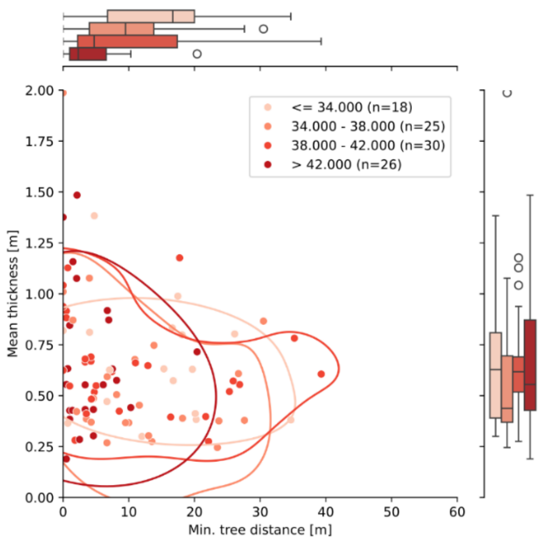

Figure 3: Analysis of the relationships between "Mean thickness"(y-axis) – "Min. tree distance"(x-axis), colored and classified by "pre-failure slope"(coloring) of the scars (≤34, 34-38, 38-42 and >42 [deg.]).

It should be noted that when the mean DoD/depth value is considered, the slope classification does show significant differences (p-value of 0.031). This is also what is expected, since the DoD value is highly dependent on the slope, as it is measured in the z-direction which results in higher measurements with higher slope values.

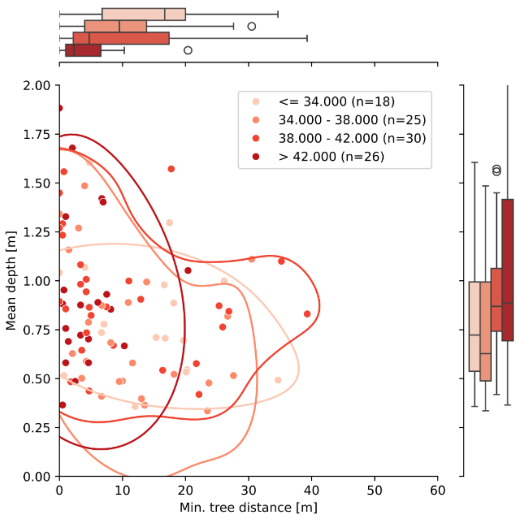

Figure 4: Analysis of the relationships between "Mean depth (DoD)"(y-axis) – "Min. tree distance"(x-axis), colored and classified by "pre-failure slope"(coloring) of the scars (≤34, 34-38, 38-42 and >42 [deg.]).

Regarding point 3: We also performed a separate analysis, comparing the landslides located outside the 2013 forest mask at locations of significant land-cover change with the landslides inside and outside the forest mask at locations without land cover change (Fig. 5). The analysis showed that the landslides located outside the 2013 forest mask with significant land-cover changes (e.g., locations of deforestation) are very similar in their characteristics compared with landslides located inside the 2013 forest mask. These similarities could indicate that root cohesion still plays an important role in slope stability after clear-cutting, previous studies have shown that the decrease in root cohesion after cutting is a gradual process (Preti, 2013; Sidle and Ochai, 2006). Another explanation could be a difference in how the hydrological properties of the soil are affected by the bio-ecological activity in the forest and land-use practices in the meadows, which could persist after the landcover has changed.

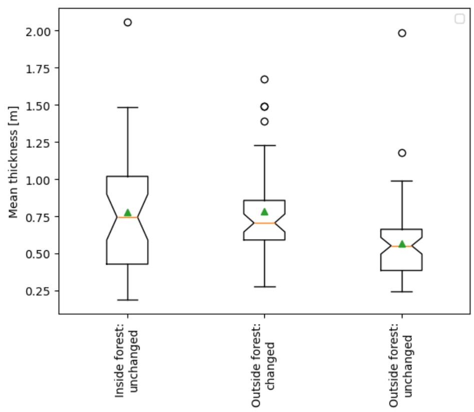

Figure 5: Comparison of the landslide characteristics inside and outside the forest, with a separate classification of the landslides outside the forest which changed land-cover class since 1995.

The results of the manuscript will be updated and only the analyses that exclude the landslides at locations of land-cover change will be used in the final manuscript. We also think that it makes more sense to use the minimum tree distance of the scar in the results and will replace the figures using the average minimum tree distance. Additionally, the analysis of the effect from the slope on the landslide thickness will be added to the results section. Lastly, the discussion will also be updated to incorporate the new discussion points addressed above.

Other Notes:

- LN 99-104: Is this a naturally occurring grassland? What is the land-use history of this area and how might that be reflected in the local vegetation patterns? Can the authors elaborate on the characteristics of the trees in the context of landslide-relevant metrics (e.g., canopy cover, rooting density and depth, etc.)?

The grasslands present in the study area are not naturally occurring, but have been extensively cultivated for a long time (> 200 years). The forest has also been cultivated for a long time. When we analysed historic orthophotos (from the 1970s to 2013) of the study area, we found that there have been some significant changes in the forest in the study area. Generally, the forest has extensively grown since the 1970s. However, there are also many locations within the forest that have been clear-cut or damaged over the years by for example windthrow events.

For future research it would be very interesting to investigate the effects of land-use and land-cover changes on the landslide characteristics in more detail. For this study, this analysis lies outside of its scope.

LN 100: How was the forest canopy cover mask created? Did the authors generate it or did it come from another source? If the authors created it, should it be introduced in the Methods, as opposed to the Study Area?

It was created during the study. The methods for its construction are given in section 3.1. The authors agree that it would normally be better to introduce the forest mask in the methods. However, since the forest mask provides important information that describes the setting of the study area, the authors decided to describe the forest mask already in the study area description.

- Ln 127-128: Should a citation be provided for “seeded region growing” algorithms?

The authors agree. In the updated manuscript a citation was given for the used tool.

- LN 148-151: Is the tree location dataset attributed with any information about tree type?

The algorithm for the forest mask construction does not provide information about the tree type. Existing general datasets tell us that the predominant species in the study area is Norway spruce. For further investigations into the effects of vegetation on landslides in the study area, tree species data would be very helpful information that could be used to analyse the effects of different root systems from different tree species. However, this analysis should also consider other important tree-specific data such as tree age, height, local growth conditions, elevation and aspect effects. Collection of this large dataset lies outside the scope of this paper, instead we have added such an analysis as outlook in the discussion section of the paper.

- LN 228: Use of "scar width" (if that is the intended meaning) may be more intuitive than use of "scar thickness," especially since "scar depth" is already used.

Scar width is not the intended meaning. The scar thickness is a "depth" estimate that is measured in the slope normal direction instead of the z-direction, as the scar depth. Since the scar depth (DoD measurement) is highly related to the slope of the terrain, the use of the scar thickness is more suited for the comparison of landslides from different terrain types.

- Results: I understand that the authors employ a semi-automatic mapping routine to reduce subjectivity in landslide delineation across the study area, but how is the transition between source and runout handled? This seems important, as landslide area is one of the metrics that the authors consider and this metric would also affect landslide volume. With sufficiently high-resolution optical imagery and lidar, a human can usually distinguish the transition from the source area to the transport zone for shallow landslides. It's unclear what the segmentation algorithm considers.

The segmentation algorithm is an additional step converting the probability output of the random forest model into a binary classification map (landslide vs. no landslide). It is more common to use the direct classification of the random forest model for the final output of the landslide detection, however, since the classification of the RF model was determined to have a high false positive rate, it was decided to further process the probability output. The goal of the developed segmentation algorithm in the detection workflow, is thus to automatically remove false positives from the detected landslides by the RF model.

In a follow-up step the segmented areas of the RF probability are used together with the convergence index of SAGA GIS to map the actual landslide scars. The assumption is made that a landslide scar (source area) can be described as a pit within the DoD dataset. The convergence index provides outlines of separate pits, which are then assumed to represent the scar boundaries of the landslides.

- LN 287-290: Suggest reduced use of "it is interesting" unless the authors are more explicit about why these statements are particularly noteworthy.

This is changed in the updated manuscript.

- LN 395-397: Can the authors be more specific about these discrepancies?

A comparison of the BFW and RF inventories shows a slight underestimation of the landslide size by the automatic mapping from the convergence index. However, this was also expected since the mapping of the BFW inventory with the seeded region growing algorithm relied more heavily on the orthophoto, which generally shows larger landslide signs than the DoD data (at least in this study area). The RF inventory is still preferred for the investigation, since the determination of the scar boundaries is more robust than the boundaries in the BFW inventory which are biased by the operator.

- LN 409-411: Why might landslide area and landslide volume be unrelated to forest location? Landslide depth (found to have a relationship) would ultimately factor into landslide volume, for example.

On average the scar area is ~50x larger than the mean thickness of the scars. In the calculation of the scar volume, the thickness thus only has a small impact on the volume of the scar and because of this, the influence of the forest location on the landslide thickness does not show up in the analysis of the scar volumes.

- LN 416-418: Can the authors elaborate as to why distance to a tree appears to have a stronger relationship than the number or density of trees surrounding a scar?

First it should be noted that this still is true with the new results. The authors suspect that the tree density metric is not able to explain the variation in scar thickness, since the value is determined at a fixed tree distance (10 m) and this ignores the variation of root cohesion in relation to the distance from the stem of the tree (Schwarz et al., 2010) and similarly also the variation over distance of other processes that are affected by the tree.

Referee #2:

The authors thank Reviewer #2 for his comments on the manuscript. Please find our answers to the comments in green text below.

Thank you for giving me the opportunity to review this paper, I enjoyed reading it.

The authors aim to increase our understanding of the effect of vegetation, specifically forest cover, on landslide processes. To achieve this, they develop a new semi-automatic algorithm that combines topographic data (point clouds and DEMs from ALS) with aerial imagery, using a machine learning approach (Random Forest) to map landslide scars after an intense rainfall event. In addition, a forest cover mask is created to distinguish between landslides triggered in forested and non-forested areas. The authors report significant differences in landslide scar characteristics between these two environments, such as greater depth, thickness, and pre-failure slope in forested areas.

As I mentioned earlier, I appreciated reading the paper. It is well-structured and pleasant to read, with clearly stated objectives. The issue of landslide under-detection in forested areas is indeed a common limitation in landslide inventories. I commend the authors' effort to use an approach that integrates different types of data, which seems appropriate to meet the objectives of the present study. Nevertheless, I suggest that the authors make some modifications and add clarity to a few points before the paper is published.

Best regards,

Thomas Bernard

Specific comments

. Abstract: Well informative. I would just add that the landslide detection workflow is based both on Lidar and the use of a random forest algorithm to make it clear that it is not just based on topographic difference.

Thank you for the suggestion. This is now added to the abstract.

Section 3.1: It would be useful to add a table with information on the lidar dataset (date of acquisition, mean point density, ground point density, vegetation points density, data availability). The points density is informative on the quality of the data you are using. It is quite important especially for the construction of the canopy cover mask, the detection of single tree position and the derivative maps (tree distance and tree density maps).

A table was added to the updated manuscript.

L 163: "(original DoD downscaled to a 0.1 m resolution)". I guess the authors choose to increase the resolution of the DoD to match the one of the orthophoto. I have two comments. First, why making this choice rather than decreasing the orthophoto to 0.5 m? In this way the interpolation is based on real information. Second, can you explain the method used to downscale the DoD? Did you produce DTMs at 0.1m from the point clouds? Maybe add this information in the text.

The 0.1m resolution was used, since the information by the orthophoto was weighted heavier in the creation of the validation inventory. This was done on purpose since it would provide a training and validation dataset that was not created with the same dataset as the dataset used for the RF model input. It should also be noted that although the initial result of the scar boundaries of the validation data were produced automatically with the seeded region growing algorithm, all polygons were also analysed and corrected in a subsequent visual analysis with the DoD data, its derivatives and the orthophoto. For the creation of the 0.1m DoD, the 0.5m DTM was resampled with a nearest neighbour algorithm, which limits the amount of new information being introduced in the input of the segmentation.

L 171-172: Are the non-landslide areas delineated automatically or manually?

The non-landslide areas were delineated manually based on the visual analysis of landslide signs on the orthophoto and DoD.

. L182: There are some technical terms such as hypertune that are unclear for non-expert or non-users of this kind of approach. Maybe add a sentence to explain a bit even if the reader can access to the reference of the method used.

This is changed in the updated manuscript.

. L 191/265/269-270: Can you explain a bit why there is an overestimation of landslide locations? Do you have an idea on the number of false negative? Especially in forested areas.

The FPR of the probability output of the random forest model (at the threshold with the best D2PC value) is 13% with a TPR of 85% (with a pixel-based analysis). Note that we preferred a lower FPR over a higher TPR, since a lower FPR attaches a higher certainty to the inventory used in the analysis. A visual inspection of the probability output shows that the overestimation is mainly related to non-topographic changes in the meadows. Due to the differences in grass height, large sections in the meadows show either positive or negative change in the DoD and the RF model is not capable of filtering this out. With regards to the false negatives, an analysis using an intersection of the final RF inventory with the BFW inventory shows that 44 out of 112 that could be detected by the model are missed (i.e., excluding BFW landslides located in the masked out area). A further inspection of these missed landslides shows that 19 of the missed landslides are very shallow and difficult to detect visually in the DoD data (mean depth < 0.5 m). An additional 6 missed landslides were also part of a complex group of landslides and their separation was handled differently in the BFW inventory than in the RF inventory, or the CIX only provided one pit location. Lastly, for some landslides the topography of the landslides was also changed after the event, following restoration works from farmers.

. L220: “.. landslides with partial cover <90% were excluded from this binary analysis.”: I understand the point, but I don’t think it should. It would be interested to see the same results in this category because if they are not different than the >90% canopy areas, then it’s difficult to claim than the forest cover is the main parameter explaining the observed differences.

There are only 19 (in the new dataset) landslides which have a partial (between 10 and 90%) cover of the forest mask, because of this it is not possible to perform a representative statistical analysis of this landslide subgroup. In our opinion the forest mask is also less suited for representing the effects of vegetation presence on slope stability, since the forest mask cannot represent the forest structure, unlike the forest structure parameters.

. L231-232: Why not using directly the point clouds to compute normal (thickness) distances to detect changes directly with M3C2. It is more appropriate than doing it from the DoD (vertical distance), especially on steep slopes (see Bernard et al., 2021).

The use of the point-cloud data for extracting landslide characteristics was out of the scope of this work. But the authors fully agree that this would lead to more accurate metrics of the landslide characteristics.

. L 323: The general convention for significance is with a p-value < 0.05. Can you please explain why do you consider these results significant with a p-value of 0.10?

The authors of course fully agree with this. The statement was also not meant to imply that the difference (in landslide thickness with the binary forest mask separation) is statistically significantly different. It was added to provide a more thorough description of the results. In the updated manuscript this section has been updated with the new results, which do show a significant difference (p-value < 0.05) in the distributions of the landslide thickness values.

. L 384: Is this underestimation compared to the BWF inventory? It is not clear. Also, can you say a few words why there is an underestimation?

Yes, this is correct there is an underestimation of the mapped scars compared to the BFW inventory. As stated in the manuscript this underestimation expresses itself both as an underestimation with regards to the number of occurred landslides and the area of the individual landslides. The workflow is not capable of capturing all of the landslides that occurred. See our previous answer with a reasoning for the missed landslides on the comment regarding: L 191/265/269-270. Besides this the final scar inventory also generally underestimates the landslide area of the individual landslides, when they are compared to the individual landslides of the BFW inventory. This was expected since the mapping of the BFW inventory with the seeded region growing algorithm relied more heavily on the orthophoto, which generally shows larger landslide signs than the DoD data (at least in this study area).

. L 385: Is the false positive rate pixel-based or polygon-based?

The mentioned FPR is pixel-based. This has been clarified in the manuscript.

Technical corrections

. Figure 1: There are too many subplots which make it difficult to read, especially for the DoD (11.c)). I suggest splitting the figure in two: 1.a) and 1.d) together and 1.b) and 1.c) in another figure (maybe located in section 3.2 or 3.3). Also, because the DoD is kind of a result by itself. Also, not being familiar with the area, I think the figure 1.a) does not give enough information on the location, maybe add main cities would help to understand the location.

The figure has been split up in the updated manuscript and the overview figure has been provided with additional information on the location of the study area.

. L 84: "focusses" to "focuses"

This is changed in the updated manuscript.

. L 121: "… 10 landslides are located within the areas with high tree density.". What's high tree density? Please give some quantification.

High tree density was meant to describe a location inside the forest mask. The authors agree that this is confusing and it is changed to "inside the forest mask" in the updated manuscript.

. L 127-128: "… using a seeded region growing algorithm". At this point the reader is wondering how this algorithm works. I suggest referring to section 3.2 here.

This is changed in the updated manuscript.

. L 151: "… by thresholding the canopy cover at 30%". I did not understand 30% of what? Please can you clarify?

This is calculated in the algorithm from LIS Pro 3D (LIS Pro 3D, 2025), as described in Eysn et al. (2012). The algorithm bases the construction of the forest mask on the convex hull area of a "tree triplet" group and the crown areas of the individual trees in this triplet group. If the total crown area of this triplet divided by the convex hull area (assumed to represent the canopy cover) is > 30%, the convex hull is kept and used in the further construction of the forest mask.

It is now clarified in the manuscript that this is the canopy cover calculated by the algorithm.

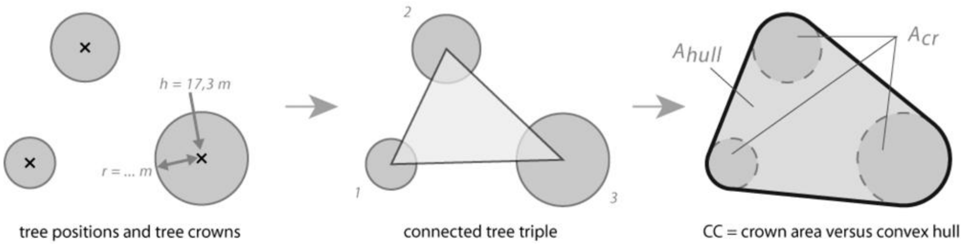

Figure 6: 'Tree triples' approach: three trees are connected for calculating the CC. The amount of CC is the relation between the area covered by crowns and the area of their convex hull (taken from Eysn et al., 2012).

. L254: "TPR" and "FPR" are used without being previously defined. Please correct

This is changed in the updated manuscript.

References used in the answers:

Eysn, L., Hollaus, M., Schadauer, K., and Pfeifer, N.: Forest Delineation Based on Airborne LIDAR Data, Remote Sensing, 4, 762–783, https://doi.org/10.3390/rs4030762, 2012.

LIS Pro 3D (2025): Point cloud processing software. Version 2024.03 [Online] https://lispro3d.com, Accessed on: Jul. 3, 2025.

Moos, C., Bebi, P., Graf, F., Mattli, J., Rickli, C., and Schwarz, M.: How does forest structure affect root reinforcement and susceptibility to shallow landslides?, Earth Surface Processes and Landforms, 41, 951–960, https://doi.org/10.1002/esp.3887, 2016.

Preti, F.: Forest protection and protection forest: Tree root degradation over hydrological shallow landslides triggering, Ecological Engineering, 61, 633–645, https://doi.org/10.1016/j.ecoleng.2012.11.009, 2013.

Schmidt, K. M., Roering, J. J., Stock, J. D., Dietrich, W. E., Montgomery, D. R., and Schaub, T.: The variability of root cohesion as an influence on shallow landslide susceptibility in the Oregon Coast Range, Can. Geotech. J., 38, 995–1024, https://doi.org/10.1139/t01-031, 2001.

Schwarz, M., Lehmann, P., and Or, D.: Quantifying lateral root reinforcement in steep slopes – from a bundle of roots to tree stands, Earth Surface Processes and Landforms, 35, 354–367, https://doi.org/10.1002/esp.1927, 2010.

Sidle, R. C. and Ochiai, H.: Natural Factors Influencing Landslides, in: Landslides: Processes, Prediction, and Land Use, American Geophysical Union (AGU), 41–119, https://doi.org/10.1002/9781118665954.ch3, 2006.