Stantec

The authors are grateful to the reviewer for his constructive comments. We'll endeavor to respond appropriately below.

Structure:

[*The manuscript includes many figures (27) and possibly some of then can be further combined and or excluded or also improved (screenshots by themselves are not very informative). I think that legend is important in the figures showing the erosion and deposition cells.*]

We agree that some of the Figures can be combined and will do so on the revised manuscript. We'll remove Figure 2 and replace it instead with the underlying algorithm.

[*Authors state that the software will be freely available for to no-profit groups. I wonder why not to release the code under an open source license (GPL as an example) in order to facilitate the re-usage and improvement, of this interesting tool, by the scientific community. I also wonder if the tool will be provided as a binary code for multiple OS (Win, MacOS, GNU/Linux)*]

Currently, we are not intending to release the code under an open source license. We are extremely interested in working with the larger scientific community and will freely provide the software for non-commercial use. The tool was written for the Windows platform and will not be provide as a binary code for multiple OS.

[*Among the different papers dealing with regional modeling of debris flow runout cited in the introduction I would suggest to introduce these other two papers that are exactly dealing with the topic:*]

Thank you for the references, we will include them in our update. The papers by Mergili and others take a decidedly different (though innovative) approach to landslide runout. We expect, based on our review, that they require additional software (UNIX, GRASS GIS) that are less common outside of academia, as well as calibration of the break criteria. While LABS also expects calibration by an expert, the methods are, in our opinion, more accessible to the user.

[*In the description of the program it is said that the software uses 5m resolution DEM. Is it not possible to run the tool on a grid having different size? If not I think that some comments on the memory requirement for running the tool on a given portion of the territory should be provided since this can hamper to use of the tool, on normal laptop or workstation, for modeling large areas.*]

The program is optimized for a 5 m DEM. Smaller or larger DEMs are likely to produce incorrect results. Comments about computer power and memory are almost certain to be obsolete at time of (or shortly after) publication. We are currently running the program on a variety of laptops and limiting processing time by breaking down the DEM to reasonable sizes (currently several hundred km^2). For slower computers there is a Region of Interest function within the program that allows the user to define an limited area for analysis.

[*Description of the program is rather short and not very detailed. As the tool is proposed to the scientific community I suggest to enlarge the description of the algorithm. As an example section 2.4 about spread is not very clear to me and the 3D Gaussian surface of figure 2 is not very informative.*]

Agreed. We will replace Figure 2 with a more detailed explanation of the spread mechanism. In particular we will add the following:

"Landslide Shape and Spread (spawning) are described by a probability density function where the mean is centered around the facing direction of an individual agent (accounting for the local slope by way of the Moore neighbors) and the standard deviation, σ, is defined by:

$$\sigma = \left(\left(\frac{m_{MAX} - m}{m_{MAX}}\right)^n * \left((\sigma_L - \sigma_S) + \sigma_S\right)\right)$$

Where: m_{MAX}=Fan Maximum Slope, m=DEM slope, n=Skew coefficient , σ_L=σ Low Slope coefficient, σ_S=Steep Slope coefficient.

These are controlled, in turn, by sliders within the program itself that cover:

- Fan Maximum Slope
- σ Steep Slopes
- σ Low Slopes
- Skew Fanning to Low Slopes"

[*I also suggest to clear if the model, after the run, alter the DEM, carving or uplifting it in correspondence of erosion and deposition areas. This would be an important tool since, when a an agent passes through a given cell, it could find the the DEM altered by an antecedent agent coming from another source and this could have an effect on the propagation of the second one. It seems the DEM is modified, from what we can read at lines 10-11 at page 6, but probably it should be made more clear.*]

You are correct, the model alters the DEM as it propagates and this in turn affects subsequent landslides. However, once the run is reset, the DEM is reverted to it's imported (original) form. This ensures that multiple runs are not affected by previous runs.

[*Some concepts are repeated. As an example page 5/6 lines 26-27/1-2 or page 7 lines 10-11*]

Right. We'll review and remove unnecessary repetition.

[*I suggest to add information to the background for the first case study. As an example it is relevant to know the size of the study area, which DEM was used and if the DEM is pre- or post-event.*]

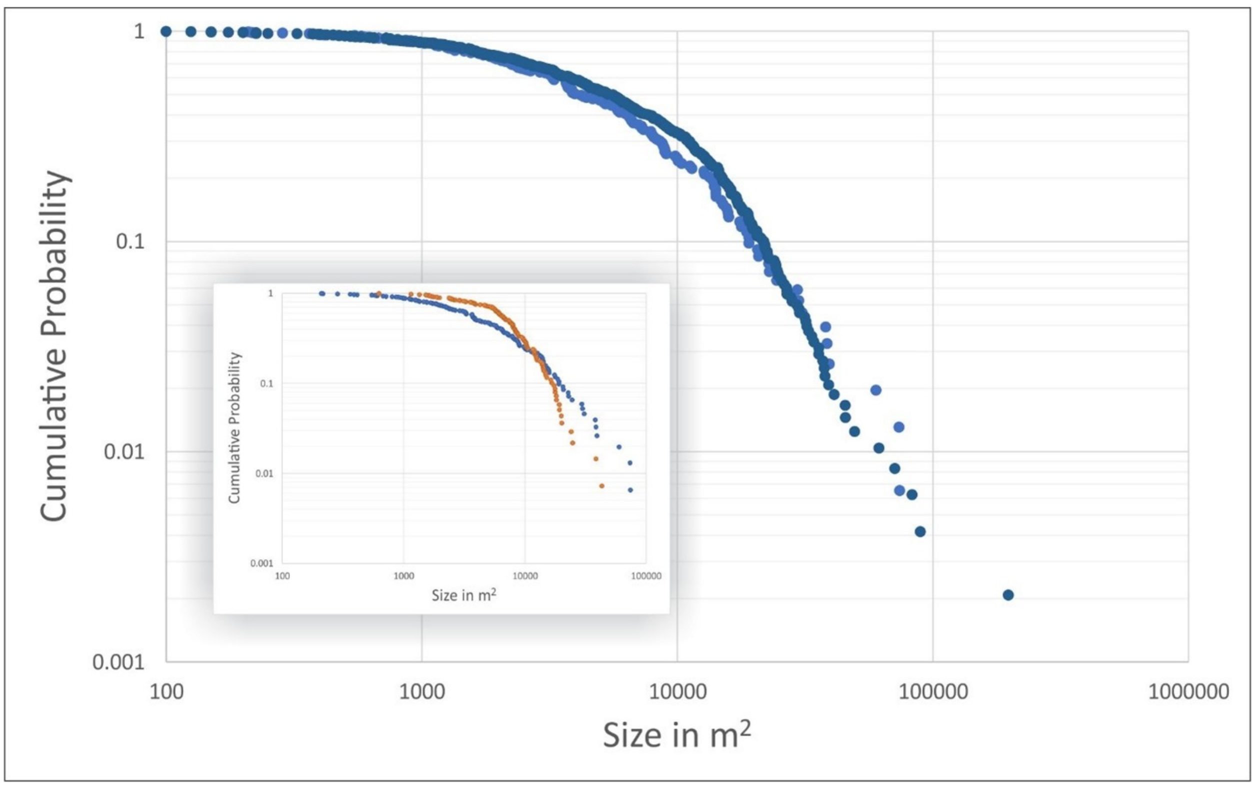
Great points. We'll provide the additional detail. For the record, the study area is 21.4 km^2, relief is 2,646 m, the DEM is 2018, and it is post event.

[*Section 3.2.2 is about calibration. However it is not clear how the model is calibrated. I would have expected that some of the parameters used by the model would have been changed to make the model match with the observed data but it seems to me that there is not such type of action. Why didn't the authors tried to tailor the model results to the ground truth? This has also to do with the model sensitivity. I really suggest to better discuss this point.*]

We agree that this section needs additional explanation.

Model calibration is completed iteratively using the controls within the program. The landslide professional runs the model per Section 3.2.2 and compares the results to mapped or historical landslides and ground-based evidence for travel distance, scour and deposition. Several methods may be employed including a visual comparison, magnitude-frequency comparison of mapped versus modeled landslides (e.g. Figure 13), and volume area relationships (e.g. Figure 7). The “Inspect” tool allows the user to examine the results pixel by pixel and the “One By One” advances individual agents through single time steps allowing for a much more detailed analysis of results.

Typically, adjustments are made to the control sliders until better results are realized. This might require several runs. Control sliders adjust the shape and spread, and the volume eroded or deposited in each timestep. Note that the volume controls are new since the manuscript was submitted. The attached figure shows the difference between a poorly calibrated result and a well calibrated result using M-F analysis.



Example of a M-F graph of modeled and mapped landslides from a well calibrated model run and an earlier poorly calibrated model run (inset).

[*In section 3.2.3 it is said that random source location are placed based on a susceptibility map. It is relevant to describe how this susceptibility map was generated. Authors also mention existing terrain polygons but, if I’m not wrong, they were not mentioned before and I don’t know what they are. It is not clear to me what they are. In the same section lines from 1 to 8 are not very clear to me. Probably experiment settings should be better explained. Again, in the same section, at line 10, authors say “once* *the historical event were calibrated..” but calibration phase was in section 3.2.2 and I’m not sure they are talking of that part of the manuscript but rather about the six storms.*]

One of the biggest challenges we had writing this paper was to provide enough information to elicit interest in the model and show that it had wide application and robust performance while, at the same time, respecting typical article length. The susceptibility map was relevant in that it provided a logical place from which to initiate future landslides, however, methods of creating such maps are well understood and, in our opinion, unnecessary to repeat here. Suffice it to say that whatever the method of susceptibility mapping one uses, the landslide professional can then choose to initiate landslides from appropriate locations within the landscape.

We will nonetheless seek to clarify the section in the updated manuscript.

Line 10 should read, “Design floods were determined by bulking the debris flood model with sediment estimated for specific storm return periods from the calibrated debris flow model using LABS (Table 4).”

[*At lines 14-16 of page 9 authors discuss a sort of susceptibility map. The citation is Palmer (2018) but the reference is “Palmer: Lake Cowichan and Youbou Slope Hazard Assessment, 2018.” that I wasn’t able to find.*]

The report is publicly available here: Natural Hazard Risk Assessments | Cowichan Valley Regional District (cvrd.ca)

[*At line 20 of page 9 authors say “The model was calibrated by simulating landslides within the study area, comparing the results to mapped and expected landslide behavior..”. I think they can improve the description of what they intend with “mapped and expected landslide behavior”.*]

LABS is both predictive and probabilistic. We are not precisely recreating an existing or historic landslide, but instead trying to credibly produce predictions of landslides that may occur on the existing surface. We think that this is a strength of the program, however, it means that calibration includes a degree of expert judgement. The professional must decide whether modeled landslides travel along realistic paths, whether the paths are similar to those of historical events as mapped or as observable in the air photographs, whether the range of deposition and erosion approximates similar events in the same region, and finally, analytically, whether or not the magnitude frequency characteristics are sufficiently similar to mapped characteristics, or justifiably different.

[*At line 24 of page 9 authors say “magnitude frequency curves that are similar to other coastal BC data sets (Figure 18) with a similar rollover and distributions” but, taken the same area value, the model data differ till 1 order of magnitude of cumulated probability with the other curves. It seems, as a consequence, that area distribution is underestimated. Why non to try to better calibrate the model parameters to improve the matching of the model outputs with the frequency size distribution of the real inventories?*]

We agree that this section could be clearer. We propose to change it to a more precise analysis:

The tangent of the slope at a given probability of occurrence [Updated figure X below] was approximately equal for both modeled and mapped landslides. We thereby interpret that the model does a good job representing variability in landslide size distribution. However, mapped landslides generally occupied about twice the area of modeled landslides.

Mapping is, in and of itself, a model. There are restrictions related to level of detail and a practical mapping scale. The mapper must make a choice between outlining landslides that are inferred to exist on steep slopes and precisely following the limited path visible among trees. In this case, the model appears to have better limited the landslide width to the actual path [New figure XX below]. Mapped landslides include areas of steep gullies and slopes that are heavily forested after the identified event. We therefore interpret that the magnitudes of the mapped landslides are conservatively inflated and that is reflected in the curve in Figure X.

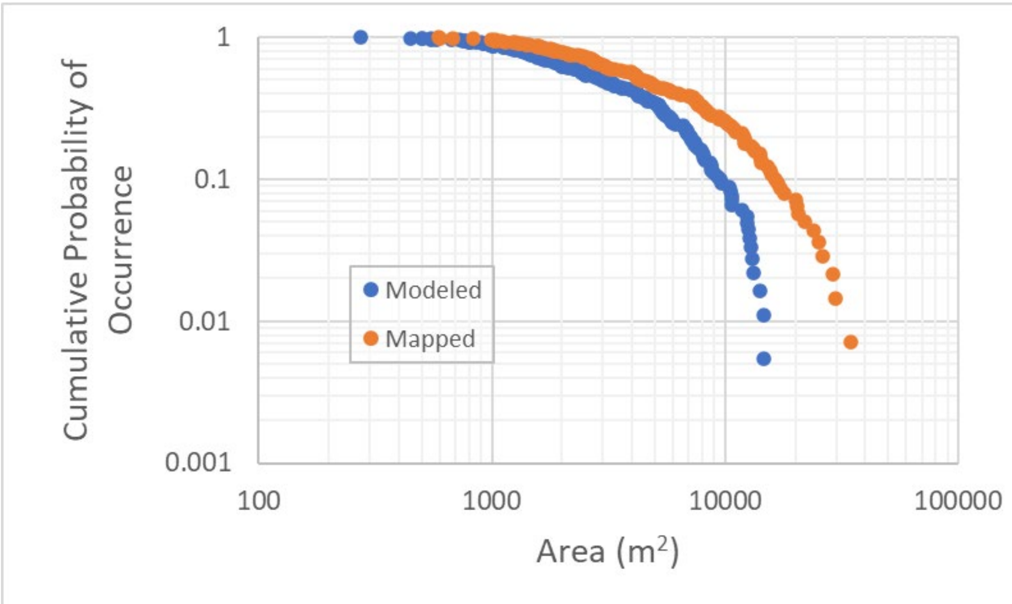


Figure X. Magnitude-Frequency comparison between Modeled (blue) and Mapped (orange) landslides on the North Shore.

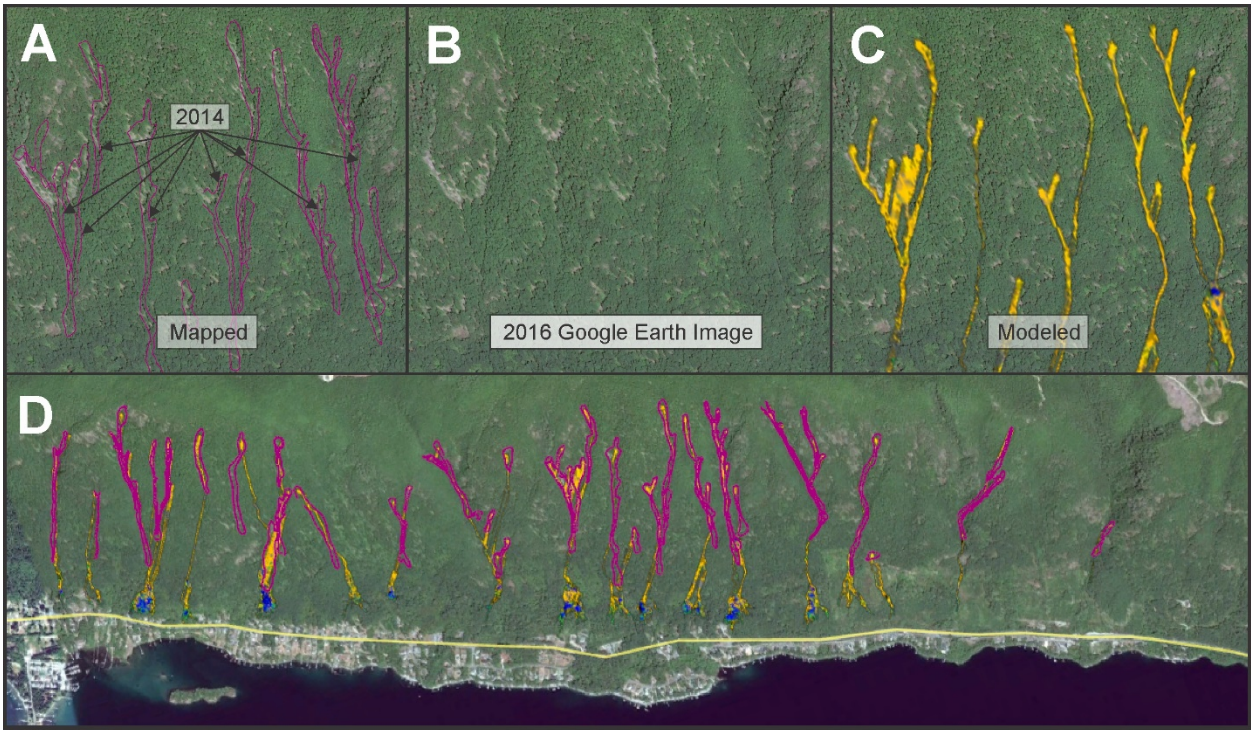


Figure XX. Mapped vs modeled landslide paths. Landslides mapped on the 2014 air photographs (A) are compared to similar modeled (C) landslides against the 2016 air photographs (shown in all instances here – imagery from Google Earth).

[*In section 3.3.2 manuscript declares that "Landslide initiation locations were created by importing randomly distributed points, a uniform distribution of points, and manually in the GIS tool within LABS". However there is no discussion about the effect of these different methods used to define the landslides initiation locations on the performance of the models. There is only a rapid comment on section 3.3.3. On section 3.3.3 authors start with: "Once tested..". Are they meaning "Once calibrated.."?*]

There was no difference in the results, except that some random/uniform points didn't result in landslides (local slope was too flat). For landslides that initiated, the results were comparable.

[*At line 9 on page 10 there is the following text: "(both random and manually selected)" . But just a paragraph above authors state: "A user-based initiation-point selection method was used for the final model runs as this method generally resulted in landslide generation somewhat more frequently than randomly or uniformly generated points that would sometimes occur on a flatter portion of the slope". Sorry but I don't understand which is the method used, at the end.*]

Same as above. Using the random or uniform distribution of initiation points meant that some agents were generated on local slopes too flat to initiate a landslide response. Manual selection simply reduced the probability that this would occur.

[*Comment on lines 12-14 of page 10 are interesting but I wonder if they can be considered conservative given the fact that the magnitude frequency curve of the modeled landslides resulted in smaller landslides respect to those observed in other similar zones*.]

It's a fair question. Under current conditions, modeled landslides traveled consistently further than mapped landslides. Fanning behavior modeled did approximate vegetation changes on the fan, but exceeded what had been observed in the last several decades of air photograph interpretation.

As per the discussion above, we interpret the difference in the reported magnitudes to be a result of mapping conservatism.

[*In section 4.3 there is a discussion about DEM resolution, that is fixed to 5m. As I said, I think that this should go together with an analysis of the memory requirement for running the tool. Having a fixed value for the resolution there are probably limitations (memory) about the maximum size of the area that can be studied using a computer*.]

Such a discussion will probably be outdated by the time this goes to press. Certainly by the following year. Computational demand is affected by the size of the area being processed at a single time; however, LABS has a ROI button that allows the user to focus analysis on smaller portions of the DEM if required.

[*Page 8, line 29: "700 mm" and "6000 mm" I suppose.. Is the -1 an error? Page 10, line 29: are authors meaning Figure 24 and 25? page 12 line 6: please remove "is"*]

Page 8 Line 29 should read "Annual precipitation varies between 700 mm and 6,000 mm…". Page 10 does refer to Figures 24 and 25. Agree to remove "is" from Page 12.