Reply to Reviewer 2

Thank you very much for your detailed comments and constructive suggestions. Here are our reply to what your have raised:

Overall comments This is a very good paper that illustrates the potential of the Random Forest approach for landslide risk mapping. The research approach is very sound and all factors that affect landslide susceptibility have been taken into consideration. The results are impressive with very high success of the prediction model. The paper is well written, well structured and in good English. Even though the research method relies on the use of high-level statistics, it can be understood by any land resource scientist with only a summary background in statistics. The methodological flowchart is very transparent.

Reply: Thank you so much for your positive comments.

Specific questions/comments: 1) I would suggest to pay attention to the following points: 2) Line 76-77: replace "the" by "important"

Reply: Word replaced

3) Line 79: replace "geological and meteorological" by "geological, soil and meteorological"

Reply: Word replaced

4) Line 84: replace "land resource" by "land cover"

Reply: Word replaced

5) Lines 85-86: move reference "Qin and Liu (2018)" directly after "ore mineralization"

Reply: It is done.

6) Line 96: "RF algorithm has been rarely applied". Not that rarely apparently, e.g. references Zhang et al. and Pourghasemi and Kerle.

Reply: It is deleted.

7) Figure 1. I would recommend a wider scale of colors, for instance from green to brown: the range of elevation is from less than 70 m to more than 1200 m. Currently the figure inadequately captures this high range.

Reply: It has been reproduced in the revision as suggested. Thanks.

8) In Figures 3c, 5b and 5c one does not see any magmatic veins, so why include it in the legend?

Reply: The distribution range of the magmatic veins in the study area is relatively small, which cannot be distinguished by the eyes in the figure.

9) Given the high range in elevation, I recommend adding a figure with the spatial distribution of either annual rainfall or of the period with the most intensive rainfall.

Reply: Thanks for comments. The figures with spatial distribution of annual rainfall and March-June rainfall were added (see new Figure 3c, 3d) in the revision.

10) I think it is necessary to indicate the process by which values of the different resource factors (such as river and stream buffers, lithological classes, fault buffers, sand percentages, etc.) were converted into risk scores. I guess by 'expert judgement', but who were the experts? The authors or land users or both groups?

Reply: Thanks for having raised this issue. Resource factors were converted into risk scores by field investigation and expert judgment. It is indicated in the revision.

11) Lines 231-232: "forest cover was assigned a value of 1-2" Is it 1 or 2, 1.5 or are there two subclasses of forest cover, one with risk score 1 and another with score 2?

Reply: Forests have a low proneness, so a value of 1 is assigned. It is modified in the revision.

12) Line 236: how are NDVI converted in risk scores?

Reply: NDVI was produced using late October and early November Landsat images, when crops are harvested and herbaceous vegetation is mostly withered but coniferous and broadleaf trees are still green. We think hence that NDVI represents forest and woodland coverage and vigor. which have . Thus, NDVI value can be directly used to reflect tree density (with developed root systems) that is resistant to landslide event, e.g., high NDVI indicates low risk of landslide given the same other conditions.

13) Areas with slopes <1-5° are considered to be 'non-risk' areas. But a slope angle of 5° is nearly a 10% slope, and that is quite substantial. In the area where I live 15% of all landslides are in the slope class 8-10%. Please confirm that the slope range 1-5° is not too wide, e.g. by noting the presence/absence of actual landslides in that slope range.

Reply: Thank you for reminding. The slope value is 1-3° instead of 1-5°. Here was a mistake. On the basis of field investigation, the "non-risk" areas were selected in valleys, plains, urban and water-bodies with low hazard proneness.

14) Could you explain what would have been the consequence of setting NT to 100 instead of 300? Figure 6 indicates that OOB error is already fairly stable at NT=100.

Reply: Figure 6 shows that when NT is 100, the OOB error is still in a relatively fluctuating state. As shown in Table 1, the Precision, Recall, F1 Index, Kappa Coefficient and Overall Accuracy against Validation Set (VS) are lower than 300 when NT is set to 100.

Table 1. Performance of the RF modeling vs training set (TS) and validation set (VS)

Item	Training Set	Validation Set	Training Set	Validation Set
Number of Trees	100		300	
Precision (%)	99.08	94.29	99.08	95.00
Recall (%)	98.18	87.42	98.18	88.67
F1 Index (%)	98.63	90.42	98.63	91.73
KC (%)	97.24	80.86	97.24	82.99
OA (%)	98.62	90.43	98.62	91.49

Technical corrections: Lines 134-137: rephrase and simplify the two sentences, as they are currently somewhat confusing. What you want to say is (1) that landslides are more likely on bare land as compared to vegetated areas, (2) slope cuts and excavations for roads and housing exacerbate the risk.

Reply: Thanks for your suggestion. It has been revised and the indication of NDVI clarified.

Line 194: replace "160-120 m" by"60-120 m".

Reply: Thanks for your modification. It is done.

Lines 261 and 353: replace "realistic" by "actual".

Reply: Thanks, it is modified as you suggested.