Authors response

Dear reviewers,

Thank you very much for the valuable comments and remarks.

Based on your suggestions, we revised the manuscript and answered open questions in our replies: AC1 and AC2.

Generally, we polished the text in nearly all sections. The manuscript got a new structure. New figures were introduced. Additionally, we provide python scripts (Zenodo code) used to generate post-processing figures in Sect. 4. In the following, you can see our point-by-point response (see also Replies AC1 and AC2).

Specific questions/comments Reviewer 1	Reply
Row 52: please indicate some examples about the codes available.	We included some GitHub references for the codes.
row 53: please indicate the name of the tools. As reader I could appreciate very much a briefly overview and description of the most relevant tools for LSA.	We could include the names of the tools. However, intending to provide a technical paper rather than a review paper, we would like to avoid going into details of existing tools.
row 53: In spite of you cited QGIS, I cannot see any QGIS tool in the reference list (only arcgis and grass). I would suggest to cite: Titti, G.; Sarretta, A. CNR-IRPI-Padova/SZ: SZ plugin, 2020. doi:10.5281/zenodo.3843276	Yes, a tool for QGIS was missing. We added the proposed reference.
Sec 2: please add a list of the main libraries that are required to use and try to cite their reference.	Agree, we included the main libraries with corresponding citations.
Tab 1: it is not clear. If the goal is to show the functions or buttons, I think could be more useful to indicate all of them in a screenshot of the main window like Fig 3 where you can put all the indications about the utilities that you want to show.	Table 1 was used because the application has three tab-frames, which means we would need at least three figures, such as Fig. 3, to indicate all functions. We solved it by introducing a more detailed schematic overview replacing Table 1 and Fig. 2.
Fig 1a: change color of landslides that are not visible.	We modified the contrast of the relief to improve visibility.
Rows 147-148: it is not clear.	We rephrased that passage.
sec 4.3: This section is too much technical, I suggest to try to reduce part of the information here reported into a graph which can help to understand the possible functions and leave the procedural steps to the manual that you can put on the code repository.	We restructured the section and introduced a new figure based on both reviews. Some technical details, in our opinion, are still needed for basic understanding. Moreover, the second reviewer wished for even more technical details in the paper. With the new Figure 1, we hope to have a satisfying trade-off.
remember to add the references of the libraries used, for example I cannot find scikit-learn reference (Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.) or GDAL or matplotlib etc…	Missing references added.
Rows 276-277: it is not clear, explain better the use of model builder	We explicitly explained the Model builder.

An overview of the study area is missing. I cannot understand the location.	The study area is now better introduced by an additional map of Pakistan with indicated study area location.
sec 5.1: I suggest to prepare a schema to better understand the procedure.	OK, but not sure if a schema with so many processing steps will not become too complex to read. We introduced a schematic workflow focusing on the described steps.
sec. 5.1: please explain better (maybe in the previous part) how the software does the crossvalidation.	We extended the passage briefly describing the cross-validation and added some additional references.
Could you add the maps resulting from the analysis?	Yes, we added the resulting LSI map.
From row 450 to 460, please add some references	OK, references added.

Reply to text comments of Reviewer 2:

Text comments were considered accordingly.

Nr.	Specific questions/comments Reviewer 2	Reply
1	In my opinion, the popularity gained by the data-driven methods is not related only to the advances in remote sensing (Line 34), but also, to the increased data availability (landslide inventories and digital layers of thematic maps) and accessibility to statistical packages.	Agree. Our formulation is probably too short here. However, the advance in remote sensing is perhaps the primary reason for the improved data acquisition of landslide inventories and better parameters such as land cover, land use, and all the derivatives from the digital elevation models, including structural analyses for lineaments. Undeniably, another domain contributing to data-driven analysis is the increasing digitalization of our world, which demands efficient data mining algorithms and tools, pushing forward the development of statistical and machine learning packages. These tools are also usable for landslide susceptibility analyses. We thought it was clear from the context of the introduction. We extended the passage to set this point in a better context.
2	I don't understand why you say that "LSAT PM's core is the weights of evidence (WoE) method" (Line 71). I understand that your suggested workflow begins with WoE method for exploration of the data and then the application of other methods for comparison. But, with this sentence it seems that WoE must be run before any other method. Is it true?	It is not necessary to run WoE before other analyses. We only recommend starting with WoE since the approach allows more insight into the data analysis than different machine learning algorithms and helps in the data exploration. WoE is the "core" because we initially started to build the application around it, applying more and more different tools. We modified the passage to make it more transparent.
3	Vector inputs are limited to the inventory? How can I manage a vector Geologic map in	Yes, now LSAT PM utilizes vector data for inventories only. As part of our

	LSAT PM? As far as I could test the software, the environmental factors can be only input as raster. In this case, you should clearly specify it. And also, if this is the case, why do we need Geology and Land Use vector maps in the test dataset?	development plan of LSAT PM, we aim to include a rasterize tool for vector inputs. Therefore, it was a strategic decision to have the vector datasets geology and land cover in the test dataset. Another point is that shapefiles Geology and Landcover can be used as input in Vector Tools serving as a mask for subsetting the inventory via Geoprocessing tools.
4	Figure 1 should be improved. At least, the letter "a" and the north arrow are not visible in Fig. 1a; Coordinates can be fixed better (latitude numbers vertically oriented and without exiting the margins); Add a location map. Also, the names of the geological formation without a short description are not useful for the reader	We improved the legibility of the figure as suggested. In addition, a table with a short description of lithostratigraphic units is introduced.
5	After Line 160 I would specify the different splitting options for the inventories (random, spatial, temporal), in case the software allows to do it.	We extended the section to emphasize the splitting options provided by the LSAT PM. Also, we introduced a new schematic graph (new Fig. 1) replacing the former Fig. 2 that better shows the procedures and data requirements.
6	In Lines 168-170 you say that "vector data are unsuitable for spatial analysis", and I don't agree. Maybe, linear and point-like vector data can be unsuitable, but a land-use vector map is fully suitable for spatial analysis, in my opinion. Please reformulate this part.	There is a specification for not usable vector data in the sentence: "…such as tectonic features or roads…". Thus, we meant only vector data without areal extent, and therefore we see no disagreement with your opinion. Of course, a vector map such as land use or geology can be used for spatial analysis. To avoid misunderstandings, we rephrased the passage as proposed.
7	In Lines 180 -182, why do you say that "Contingency analysis is the only tool in the tool domain Raster data"? According to Tab. 1 and Fig. 3, there are other tools (Euclidean distance, Combine …). I would reformulate the complete paragraph starting with something like "The contingency analysis tool helps to explore ..."	Contingency analysis is the only tool in the Raster data tool, which produces an output file (table) in the folder statistics. Other tools in the tool domain Raster data do not. That is what we mean in this passage. We rephrased the passage to make it more clear.
8	In section 4, specify and describe better the data requirements and outputs obtained at each step (contingency table, result tables). Which specific information they contain?	We restructured section 4 (now section 3). The new figure 1 better represents the workflow addressing the data requirements.
9	I would remove Lines 244-245, because at this point the model builder is not introduced yet, and they confuse a little bit.	Agree. Passage removed.
10	In section 4.5 the explanation about the sampling error assessment needs more details. Provide details on how it can be done using LSAT PM.	We extended the passage to briefly describe the procedure. Details can be found in cited papers. Details on how it can

		be done in LSAT PM are introduced in section 5 “Application to the test data”.
11	When is it used the validation sample? After having tested the software, I understood that in the Model Builder module you have the option to generate the ROC curve respect to the desired inventory partition (training, test or even a group of subsamples). This is a crucial step of the evaluation that is not clearly explained in the manuscript. Please, improve this part.	Agree this functionality was poorly mentioned in the manuscript, and we improved this.
12	The Zoning module is used to reclassify the susceptibility maps in few, and more understandable, susceptibility classes. According to your description in section 4.6, it seems that you follow the approach of Chung and Fabbri (2003). However, this approach is not based in ROC curve, but in the prediction and success rate curves, which are completely different things. I believe that if you want to set the classes in a way that you can ensure the proportions of landslide areas that should fall within the specific zone, then a prediction rate curve has to be used, and not a ROC curve. I did some tests and I realized that the curve prepared with the Model Builder and the curve prepared by the Zoning module are identical. In my opinion, this has to be fixed before publishing the software. In addition, if you suggest some classification thresholds by default (50; 30; 15; 4; 1), you should explain more in detail the implications of this values in the interpretation of the maps. Because, the suitability of such thresholds can be discussed.	Yes, we use the ROC curve to specify the classification (as explicitly specified in line 279). The classification is not affected since both ROC and success/prediction have identical y-axis (true-positive rate = cumulative landslide areas). This is the only input we allow for establishing the thresholds. The y-value is directly linked to the ranked LSI index used for classification. But it is correct that there are differences between success/prediction rate and ROC curve. They differ in the representation of the x-axis, which does not include landslide areas in the ROC curve. Thus, the x-axis in the ROC curve is only an approximation of the total cumulative study area. This approximation is sufficiently accurate if the landslide areas are small compared to the total study area (e.g., generally when working with point data). For larger landslide areas e.g., exhibiting several percent of the total study area, the class proportion values obtained directly from the ROC graph are inaccurate. However, unique conditions' ranking and corresponding LSI thresholds are not affected. In the attribute table of the output raster, correct class area values are now directly estimated from raster statistics (there was indeed a bug, thank you for opening this discussion point). We agree that the labels on the graph in the Zoning GUI may cause misunderstandings and adjusted them, making it more evident that the used metrics are different from the success/prediction rate.
13	The current section 5 should be reorganized. I would include the test data description in this section. Then, it should be just a section where the potential of the tools explained in	We restructured the manuscript accordingly.

	section 4 are illustrated. The procedures to build the models should be described very briefly (maybe using tables or conceptual plots/figures) and making reference to the section 4, where more detailed explanations can be found. In short, this section should give examples of (i) what we can get as outputs and (ii) how we should interpret them. In this regard, I believe that some interesting outputs are missing, such as susceptibility maps or variables evaluation reports and contingency tables.	
14	In section 6, which is the difference between hybrid model and model ensemble? Did you perform hybrid models in section 5?	A hybrid model is a model that includes more than one algorithm in the generation process inside one model. E.g., we use WoE for all categorical data and LR for continuous data. In the model builder, the differently assessed parameters are included in one additive model. An ensemble would be a combination of two homogeneous models, which were consistently independently prepared based on the corresponding methodology, e.g., WoE model, which tackles categorical and continuous datasets (through classification), and an LR model that is doing the same. Or LR and ANN for a multivariate ensemble. We did not perform an ensemble or hybrid model in the presented example. Your comment shows us that the general differences between these two approaches seem unclear; therefore, we introduced a passage explaining the procedure, emphasizing how it can be done in LSAT PM.
15	In general, I find the conclusion section a little bit incomplete. Before going through the future implementations that are planned, I think that a real recap summary is missing. Something more detailed than only three lines.	We extended the conclusion as proposed.