

LAND-SE: a software for landslide statistically-based susceptibility zonation, Version 1.0

Mauro Rossi and Paola Reichenbach

Submitted to Geoscientific Model Development

General comments

The authors of this manuscript are presenting a very interesting software, or R-script, which allows the user to perform landslide susceptibility modeling and a detailed assessment of the quality of the model in terms of model performance and standard error of the model output. This can be done by a variety of graphs and maps in a tabular and spatial visualization, automatically generated by the script. While most of the presented script was published before by Rossi et al. (2010), the authors clearly indicate the alterations and optimizations that have been performed on this original script since the year 2010. The possibility to include new models (like regression trees) is a clear step forward and novel to the old version from 2010. Furthermore, they implemented the more stable “glm” function for the logistic regression modeling, which is widely used in statistical landslide susceptibility modeling and in ecological modeling.

Landslide susceptibility modeling is performed worldwide more and more often to provide local communities with spatial information on where the occurrence of a landslide is more probable and where people have to take precautions when outlining new housing areas or when building new houses. The authors correctly state, that in many cases the susceptibility modeling is done very simply and often the limitations of the models themselves are not reported. However, the range of model performance and the general error of the prediction as presented by the standard deviation of the modeled probability is often not reported on. With this tool practitioners, or rather fellow scientists, can perform landslide susceptibility modeling and have a detailed look on model performance and uncertainties. However, as far as I understood from the manuscript, the effect of repeatedly drawing different samples for training and testing the model is only considered for the model uncertainty, but not for the effect on the model performance measures as reported on by other authors in the field. Please see the specific comments for more details on that.

The presentation of the manuscript and software is sound, however some minor English spelling and Grammar errors were identified which make the language sometimes less fluent or precise. Furthermore, the manuscript can still be improved by adding some more details on the methods and assumptions included in the software (e.g. on the sampling and partitioning into training and test sample procedure). Although the authors state clearly that e.g. the discussion of the advantages and disadvantages of the one or the other model performance measure was beyond the scope of the paper, a general discussion of the limitations of the models or the presented software and its results is missing and should be included as this is also demanded in the guidelines of this journal. Additionally, the audience of the software is unclear as the amount of information on limitations and proper usage might vary significantly if the software is aimed to be used by informed, modeling experienced scientists or less modeling experienced practitioners (please see the more detailed thoughts in the specific comments section).

Given these general comments and the following specific comments I would like to suggest minor to major revisions for this scientifically valuable manuscript.

Specific comments

I would like to suggest some restructuring of the manuscript as some essential information, such as how the sampling of presence data was performed for grid cells or terrain units, is only presented in the applications section. This is crucial information on the model which should be presented with the model in section 2. Please consider including more detail on this in the input data preparation section. E.g. at lines 109-112: Please clarify how exactly a landslide is represented in the model. In a later section it says that the entire landslide polygon represented in pixel of the proper resolution is included in the modeling. Given that in literature this is treated with different option I was wondering if this the only option in LAND-SE or if the user can choose how the landslides should be represented. Other authors such as Atkinson et al. (1998), Atkinson and Massari (2011) and Van Den Eeckhaut et al. (2006) report this step with different sampling designs in their research. Please clarify why you chose to include the entire polygon. A valuable source for discussing this might be the rather recent paper of Regmi et al. (2014) regarding the effects of which information on the landslides is included as presence data, on the modeling results, maybe for the discussion of the limitations of the software. Regmi et al. (2014) found that it makes a big difference which information of the landslide is used for the modeling, therefore it might also affect the models which are based on pixel as terrain units in this manuscript. Please consider including this in the discussion of the limitations of the model results.

Another area of interest is the sampling of training and test data. Here the question arises from reading the manuscript if the sampling of training and test data was done within the R-script or if this is something the user has to prepare beforehand? From the User Manual I read that the user has to perform that subsampling before. However, with that the repeatability of the model is at risk. Is there an option to include this in the model (e.g. similarly to the bootstrapping?)? Please provide some more details on this in the manuscript.

While the very often in literature suggested and very advanced possibility to create training and test samples randomly, spatially and temporally is implemented in the software, it is striking, that the sampling is only performed once for the model fitting and evaluation. I understand from reading the manuscript, that with the bootstrapping only the uncertainty or variation of the model in terms of the mean predicted probability and its standard deviation was assessed by testing multiple models. However, it is unclear how often the model should be run at a minimum or maximum to achieve reliable results and why this was not used to compute the range in the AUROC values and other model performance estimates as well. In my understanding, using the bootstrapping to compute repeated spatial or random training and test subsamples and therefore multiple performance measures, would be the same as repeated spatial or non-spatial cross validation as often mentioned in recent literature. This does not seem to portray the state of the art in this field as recently multiple authors have performed repeated random or spatial subsampling for assessing the model performance in terms of AUROC, spatial transferability and thematic consistency and have shown that with the sample, the performance measures change distinctly (e.g. Goetz et al., 2015; Heckmann et al., 2014; Petschko et al., 2014; von Ruetten et al., 2011; Steger et al., 2016). I would like to suggest to include this into the model or address why it was not done. Furthermore, the word uncertainty is used rather generally. Please make sure to be specific which type of uncertainty (model form uncertainty, uncertainty from the input data, etc.) is analyzed.

If I understood correctly the examples for the landslide susceptibility modeling are originating from two published studies from Reichenbach et al. 2014 and 2015. However, for most of chapter 3 this

stays rather unclear. Please be more out front about this fact if my assumption applies by referring to the studies at the beginning of the chapter.

The modeling of the landslide susceptibility scenarios depending on the land cover changes is very interesting. However, it stays unclear for the reader how the land cover scenarios were computed (e.g. a regression model like CLUE-V or other ways). Please consider inserting some information on this here by referring to the original study it was performed in.

The final remarks are very similar to the abstract and particularly to the introduction of the submitted manuscript. I would like to suggest to rewrite this section to give a more critical view or discussion on the software, its limitations and proper scale of application.

Throughout the manuscript I was wondering who the target audience for the software is. Please specify that somewhere in the manuscript. Depending on the target audience I was wondering if the user manual could contain some help on how to interpret the results of this software (e.g. the value range of an AUROC value and its meaning for the model performance, or the susceptibility classes). While reading I was also wondering if the landslide susceptibility classes are provided by the software, or if this is something that the user can choose. You see from my questions that I am getting very excited about the software as it could help many users worldwide to enhance their understanding for the local landslide susceptibility. Therefore, I would be very happy if you could address some of my questions in the manuscript. The supplementary material is well prepared and will aid any user of any level to run the susceptibility model.

Technical corrections

Text

The title seems appropriate for the paper. However, given the fact, that the submitted software is an optimized version of a script published by the authors in 2010, I wonder if it should be given a different version name (e.g. 1.1 or 2.0). Furthermore, I suggest a change in the sentence structure of the title to: LAND-SE: a software for statistically-based landslide susceptibility zonation, Version X.Y. This structure change seems more logical to me from an English Grammar point of view. However, I am no English native speaker, which is why I would like to suggest a thorough proof read of the entire manuscript by an English native speaker. In this section I will indicate some spelling and grammar errors as far as I noticed them. But this list might not be complete or correct! In the following I will give the sentence as it was in the manuscript with underlined parts that I would propose to add or change in the text.

Line 21: "... with additional models, evaluation tools or output types." Please delete the "s" at evaluation.

Line 22: "..., explains input and output and illustrates specific applications with maps and graphs." Please consider including the "and".

Line 32: "... since the early 1980." Maybe including the "19" would make sense to be more accurate.

Line 33: "... using different partitioning of the territory as mapping units, analysis of landslide inventories,..." Please consider changing from "partition" to "partitioning".

Line 39: "Malamud and his co-authors grouped them in 20 classes,..." Please include "and his".

Line 41: "According to them the relevant number of statistical models..." Please consider including "According to them".

Line 45,46: Please consider changing this part to: "... comprehensive assessment of the model performance, the prediction skill evaluations and..."

Line 49: "Susceptibility Evaluation), a software developed to prepare..." Please consider inserting the "a".

Line 50: "... with specific functions focused on result evaluation..." Please consider changing accordingly.

Line 53: "...evaluation tools or output types." Please consider changing accordingly.

Line 55: "... explains input and output, illustrates them with maps and graphs..." Please consider changing accordingly.

Lines 56 to 58: Please rephrase this sentence as it is very difficult to understand.

Line 59: "... test area to demonstrate the range of applications and different outputs of LAND-SE." Please consider shortening and simplifying the sentence accordingly.

Line 63: Please exchange "ancillary" for "supplemental"

Line 66: "LAND-SE, a software ..." Please consider changing accordingly.

Line 69: "... and combine different statistical susceptibility modelling methods, evaluate ..." Please consider changing accordingly.

Line 77: "datasets" instead of "dataset" Please consider changing accordingly.

Line 84 and 89: "Input data preparation" instead of "data input preparation" Please consider changing accordingly.

Line 145-146: "Model outputs" instead of "Models output" Please consider changing accordingly.

Line 177: Please consider spelling out the word software here and in any future occasion instead of using the abbreviation SW as it is easier to read and understand for the reader.

Line 198: "The elevation ranges from the sea level to about 500m and the terrain gradient ranges from 0° to 81°". Please consider the changes in this sentence for the manuscript.

Line 218: "..., landslide susceptibility zonation was prepared ..." Please consider changing accordingly.

Line 222: Please include the software it was written for. I assume GRASS GIS? Or a GRASS GIS tool within QGIS?

Line 252-253: "This application simulates LS zonation for a large territory, where landslide information is spotted and does not cover the ..." Please consider changing accordingly.

Line 261: Please consider using the word "transferability" instead of "exportability" as used by von Ruetten et al. 2011 and Petschko et al. 2014.

Line 262-263: Please use the singular: "landslide information"

Line 277: Please rephrase this sentence to be more specific on which loss of performance is commented on there.

Line 288: Please rephrase the sentence to be more specific on which results you are referring to here. I assume the resulting equation to describe the statistical relationship (e.g. the regression equation with intercept and coefficients)?

Figures

The figures are generally done beautifully and are very informative. I have only one minor remark. Please consider including a sentence of reference at the figure caption of figure one regarding the scale and geographical location of the study area. Figure 1 is the only figure that includes the coordinates around the map box. Please prepare the reader that this information is eliminated in the following figures but always stays the same for all figures. Additionally, a small figure included in Figure 1 showing the location of the study area within Italy or Sicily would be of high interest.

References

Atkinson, P., Jiskoot, H., Massari, R. and Murray, T.: Generalized linear modelling in geomorphology, *Earth Surf. Process. Landf.*, 23(13), 1185–1195, 1998.

Atkinson, P. M. and Massari, R.: Autologistic modelling of susceptibility to landsliding in the central apennines, Italy, *Geomorphology*, In Press, Accepted Manuscript, doi:doi: DOI: 10.1016/j.geomorph.2011.02.001, 2011.

Goetz, J. N., Brenning, A., Petschko, H. and Leopold, P.: Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling, *Comput. Geosci.*, 81, 1–11, doi:10.1016/j.cageo.2015.04.007, 2015.

Heckmann, T., Gegg, K., Gegg, A. and Becht, M.: Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows, *Nat. Hazards Earth Syst. Sci.*, 14(2), 259–278, doi:10.5194/nhess-14-259-2014, 2014.

Petschko, H., Brenning, A., Bell, R., Goetz, J. and Glade, T.: Assessing the quality of landslide susceptibility maps – case study Lower Austria, *Nat Hazards Earth Syst Sci*, 14(1), 95–118, doi:10.5194/nhess-14-95-2014, 2014.

Regmi, N. R., Giardino, J. R., McDonald, E. V. and Vitek, J. D.: A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA, *Landslides*, 11(2), 247–262, doi:10.1007/s10346-012-0380-2, 2014.

von Ruetten, J., Papritz, A., Lehmann, P., Rickli, C. and Or, D.: Spatial statistical modeling of shallow landslides—Validating predictions for different landslide inventories and rainfall events, *Geomorphology*, 133(1–2), 11–22, doi:10.1016/j.geomorph.2011.06.010, 2011.

Steger, S., Brenning, A., Bell, R., Petschko, H. and Glade, T.: Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps, *Geomorphology*, 262, 8–23, doi:10.1016/j.geomorph.2016.03.015, 2016.

Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G. and Vandekerckhove, L.: Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium), *Geomorphology*, 76(3–4), 392–410, 2006.