Natural Hazards
and Earth System
Sciences

Open Access

EGU

Discussions

Interactive comment on "Prediction of rainfall induced landslide movements by artificial neural networks" *by* Janko Logar et al.

Janko Logar et al.

janko.logar@fgg.uni-lj.si

Received and published: 13 September 2017

The authors appreciate the comments and suggestions for the improvement of the paper provided by Luigi Lombardo. The detailed answers to specific comments follow below.

1. This sentence is not clear (From this is not surprising to consequences). Could you please rephrase it?

Comment accepted.

2. There are 8 citations in landslide susceptibility. Shouldn't they be pruned down to 3?

The authors' intent was to emphasize that a lot of research has been done lately in the field of introducing the machine learning procedures (artificial intelligence) into the field of mapping landslide prone areas and facilitate the interested reader to find some relevant papers. However, we will reduce the number of references if so instructed by the editor.

3. Same as before, aren't they too many? I would write down 3 citations for the three topics you are mentioning.

Similar as above.

4. For landslide movements there are even 15 references. I think this is way too much.

Similar as above.

5. The actual description of the idea behind this paper is just 5 lines long whereas the remaining introductory part accounts for 63 lines. I would balance a bit more the ratio between the two by extending the part related to your contribution. For example, what are the research questions behind it? What would you expect to find (this may be a futile question at this stage but it just to give you an idea of how I believe the text could be improved). The impression that the reader has right now, is of a very long introduction, full of references at times and full of details in other parts and when it comes to YOUR work, then it suddenly runs fast, giving almost no description at all.

Comment accepted. We will better explain the purpose of our work. Within the introduction the authors wanted to give an overview of extensive work that has been done recently in this field but also to show there are still opened questions and options for the use of data acquired from monitored landslides. We were pleased to convince ourselves that similar learning-predicting procedures can give good results for two totally different landslide phenomena and wanted to share this experience with others.

6. I assume "internet sites" can be removed from the sentence without changing its meaning.

Comment accepted.

7. Wouldn't it be better to declare and then use an acronym throughout the manuscript? I believe using MLFFN would work and facilitate the reader.

Comment accepted.

8. I would delete partially retrogressively. If you use "mostly progressively" afterwards, it should go without saying it that the other component should be retrogressive.

Comment accepted.

9. I would remove "to protect" and use "close to..". As the sentence is formulated right now, it looks like the rock formation is "conveniently located" there ... to protect the village.

In fact, the rock outcrop protected the village. Once the earthflow front reached the limestone rock it practically stopped and the earthflow never reached the village.

10. Figures instead of Figs

Comment accepted.

11. I would like to say beforehand that I have not used any ANN in my studies, so my view may well by biased or I may miss something. However, what I know, is that all statistical models would suffer from multicollinearity issues in the data. Multicollinearity simply indicates that variables in a multivariate model are mutually correlated, hindering the application of the specific model. The first thing that came to my mind looking at the input data you use is that the three different cumulated rainfall may be correlated. I have looked around if ANN suffer from the same issue and found the article: "Comparison of regression analysis, Artificial Neural Network and genetic programming in Handling the multicollinearity problem" written by Garg and Tai (2012) where they tackle this issue by applying a PCA prior to the ANN. I am not saying what you did is wrong, nor that the analysis should be re-run. However, it could be a nice improvement if you calculate

C3

the Pearson's correlation coefficients for all the combinations in your input parameters and add a table so that the reader can get an idea of the structure of the data set.

The reasoning behind the use of three different cumulated rainfall is the following: Most rain induced landslides are triggered by intensive rainfall event preceded by longer rainy period. Hence, also later changes in the rate of landslide movements should be sensitive to the same combination of influencing factors. We will consider your suggestion during the preparation of revised paper.

12. Can you explain the reason why you choose a different number of neurons between the two different sites? For the Macesnik landslide you mentioned 25 neurons.

It is generally recognized that ANNs with low number of hidden neurons will generalize well but may fail to reproduce the training data set. On the other side, ANNs with high number of hidden neurons (and/or ANNs with high number of hidden layers of neurons) can reproduce well even the complex training data set but will normally perform poorly with new data sets (overfit). Therefore, the optimum architecture and number of neurons in ANN is problem specific. There are no general rules for the optimal ANN geometry (i.e. number of hidden layers and number of neurons). Therefore, the input data is usually split into learning and validation sets, the first used to establish the ANN, the later used to verify which ANN is better. In our study we used a number of different ANNs having different number of hidden neurons. Finally, only the results for ANN which gave the best results for each individual case are shown in the paper.

C4