

Interactive comment on “Rainfall feature extraction using cluster analysis and its application on displacement prediction for a cleavage-parallel landslide in the Three-Gorges Reservoir area” by Y. Liu and L. Liu

Anonymous Referee #2

Received and published: 25 March 2016

The manuscript deals with the effect of rainfall and its role on landslide deformation and failure. The authors carried out a feature extraction method for a rainfall data set and it was categorized by a cluster analysis. Rainfall indexes were computed for rainfall characteristics such as quantity, duration, and the number of raining days in a given period of time. The results were later applied and validated to a data set acquired at a cleavage-parallel landslide in the Three-Gorges Reservoir area. The landslide displacement prediction using neural networks for the rainfall input in the form of raw data, monthly rainfall, and feature extracted rainfall were benchmarked. The authors concluded that using the feature extracted rainfall method is best at predicting landslide

C1

displacement compared to the other methods and at the same time the computational stress has been reduced significantly.

Although the topic is very interesting from a scientific and practical point of view, the manuscript presents some limitations, conceptual mistakes, technical errors and is sometimes confusing to read. Consequently, it is not suitable for in the present form. The paper must undergo major revisions for publication in NHES.

The authors are strongly encouraged to review the paper in accordance to the high international standards of the NHES Journal.

General comments:

- The authors should re-organize the paper to have a coherent scheme regarding the presentation of the work carried out. At the moment the manuscript contains a lot of relevant information but it is scattered and spread around the paper in a disorganized manner. The authors are encouraged to consolidate this information inside the relevant sections of the manuscript and to avoid unnecessary repetitions.
- The manuscript lacks the relevant references in the topic and in addition only 13 references are cited inside the document from the 44 listed in the reference list. It is highly recommended that the authors should carefully look into this.
- Regarding also references, the paper is lacking of an analysis of the important results and issues raised by other studies, in particular in context of the submitted paper. Discussion of the results obtained in the submitted manuscript should be made by comparing qualitatively and if possible quantitatively with the results obtained in referenced studies.
- Basic descriptions and concepts are not defined inside the manuscript, such as: cleavage-parallel landslide and BP neural networks (to name a few).
- It is recommended that the authors revise the manuscript all over again and find the suitable words, phrases, technical terms and definitions in proper English. It becomes

C2

even more critical when the authors pretend to describe the methodology. Detailed comments:

- The abstract should be improved. Some statements like the one in Line 14 and Line 16 are misleading and should be rephrased.
- The introduction is lacking relevant references. This introductory part should be rearranged in a way that the references are supporting the stated comments (i.e. Line 31).
- The second paragraph of the Introduction should be fully rephrased and a better summary of the past studies should be carried out by the authors.
- In Line 33, the authors should explain in detail how this is difficult (with supporting references) and how their method is an improvement for this.
- In Line 52, Kurtz et al. 2014 is incorrectly referenced in the manuscript. This work is not relevant to this paper and is based on other approach for feature extraction. The authors should use another reference or delete this one.
- In Line 55, the authors should use the correct references to support the sentence.
- Line 70 should describe briefly what a BP neural network is.
- In Line 74, a description of this type of landslide should be included.
- The Methodology section should be fixed and improved. A better description should be included in order to make the reader understand better the approach used.
- Line 85 should include references.
- Line 89 to Line 92, the authors should explain in detail the reason to use an average daily evaporation? Is it because of lack of data or is it a common practice?
- Line 94 to Line 96, the authors should rephrase the statement and add the relevant reference.

C3

- Line 109 to 111 should explain, why the authors use this approach and why.
- Line 121 to 124 needs to be fixed and rephrased. It is not understandable, what the authors mean with high cohesion and low coupling in this context.
- In line 128, the authors mention, that the K-means algorithm is the most used clustering algorithm. In what context and explain the purposes.
- Line 128 to 135 needs relevant references
- Line 136. Also the sample selection affects the final results. The authors should elaborate in this respect also.
- Line 148: What does it means that it is shaped like stairs to the Yangtze River?
- Line 153 to 158: the authors should explain in detail, why are they using that control point and why is that significant. They also should include the figure of the profile and where the point is located, inside figure 1.
- Line 158: the authors should reflect on adding a new point or several others as the landslide movement is not uniform and it will add value to the validation. This will definitely have a big influence on the final results.
- Line 160: the authors should include the Zigui County in Hubei Province inside a map to understand the actual location and to understand if this will affect the final results as rainfall is spatially correlated.
- Line 171, in figure 1, the letters should go first, then the legend.
- Line 175. Using average interpolation for a long period can be misleading as the warmer days without rain will tend to influence the sample. Can the authors elaborate why not only using the evaporation related to the rainy days?
- Figure 1: the red square in the smaller map is barely visible.
- Line 196: the authors should explain how this coefficient is derived.

C4

- Line 203: the authors should explain how this optimal point was achieved.
- Line 190 to 203: It is unclear inside the manuscript, why the aforementioned coefficients were used and its relevance to the results. Can similar results be achieved using other techniques of normalization?
- Figure 3: Check on the r-axis. Point C3 is out of the range.
- Line 211: explain briefly the principle of maximum dissimilarity.
- Line 214 to 219: The authors should use the relevant references.
- Line 219: What is the stability criterion?
- Line 228: The authors should explain how they were coming to this conclusion.
- Line 248 to 249: Use relevant reference.
- Figure 5 and 6: Both daily forecasts and cluster analysis seem to give reasonable results. Is there a better way to show the improvement of the cluster analysis (a success rate approach)?
- Figure 5 and 6: One of the main advantages of the cluster analysis is based on the improvement of the computational effort. Can the authors further elaborate; is this improvement really worth it (as there will be also computation involved on the K-means algorithm)?
- The authors should discuss in the conclusion section, if this approach can be improved by selecting other type of clusters. Is there a possibility to perform a sensitivity analysis based on this?
- The conclusions are mostly a repetition of the text inside the manuscript. The conclusion section should be fixed and rephrased.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., doi:10.5194/nhess-2015-320, 2016.