Nat. Hazards Earth Syst. Sci. Discuss., 1, C91–C96, 2013
www.nat-hazards-earth-syst-sci-discuss.net/1/C91/2013/


Natural Hazards and Earth System Sciences
Discussions



# *Interactive comment on* "Application of a hybrid model of neural networks and genetic algorithms to evaluate landslide susceptibility" *by* H. B. Wang et al.

**Anonymous Referee #2**

Received and published: 12 April 2013

General Comments: This manuscript applies a combined genetic algorithm (GA)-back propagation neural network (BPNN) model to access the landslide susceptibility for Changshou Valley, China. Based on slope unit analysis, six environment variables are selected for model development. According to the model validation, high accuracies can be obtained to support the applicability of the proposed model. The topic is apparently related to the interest of NHESS; however, I suggest the body of the manuscript may not meet the standard of the journal. First, the design of experiment obviously cannot provide enough evidential results to support the advantage of the application; while it should be the main of the study (I will address comments on this in the Specific



Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

Comment). Second, authors miss essential explanations and definitions for the study area, the characteristics of landslides, the principle of variable selection, the setting of model parameters, the assessment of model performance, and they even miss discussions on their results. In my opinion, this manuscript mainly shows the practice of the proposed model instead of exploring scientific findings new or important to landslide study.

Specific Comments: Does the paper address relevant scientific and/or technical questions within the scope of NHESS? From the title it is, but the body of the manuscript may not meet the standard of NHESS. Does the paper present new data and/or novel concepts, ideas, tools, methods or results? Generally speaking, it's not new to use genetic algorithms to optimize ANNs for landslide susceptibility study. The idea of using this hybrid model addressed by the authors in the manuscript (page 356 line 6 – page 357 line 11; page 365, line 6-line 19) is that the model seems to be able to achieve a better performance for a complex-landsliding environment with large scale, compared to other linear/non-linear models in terms of accuracy and computation efficiency. However, the design of experiment in the study does not support the idea: the study area only contains 39 landslides in a small valley (? actually the authors didn't document the size of the study area), and geological condition is almost the same (page 364, line 5-6). I don't think it's an appropriate case to examine the applicability/advantage of the model. In fact, the main scope of the study is not clearly defined. Are these up to international standards? No. Are the scientific methods and assumptions valid and outlined clearly? The method and its benefit to landslide susceptibility analysis should be valid, but they are not clearly described (in section 3 and section 4). For example, among the six variables selected for the model practice, two variables, distance to river and human activities, are not explained at all. As said, I don't think the authors clearly define their main objective of the study: it looks like that to practice the model (and to achieve a high accuracy) is the main task of the study. The authors should carefully report all advantages and disadvantages of the model, and also discuss its limitations. Are the results sufficient to support the interpretations and the conclusions? No. The results



NHESSD
1, C91–C96, 2013

can be divided into four parts. The first part is about the background information which is redundant (page 364, line 2-9). The second part is associated with the frequency analysis of landslide occurrence versus each variable considered in the model (page 364, line 10-18). Plots (Figure 15-18) show the distributions without comparing them with background distributions. For example, the study concludes that 38% and 56% of landslides are toward the southeast and southwest. From Fig 7, however, the two aspects are dominant in the study area so that the significance of the frequency analysis is doubted. And again, the distance to river and human activities are not shown here. In the third part, only accuracy numbers are reported. Authors didn't comment their results at all (page 364, line 19- page 365 line 2). The fourth part (page 365, line 3-19) is just a description to highlight the advantage of the GA-BPNN techniques, without any supports conducted from their results. To sum up, the section of "Results and discussions" is poorly organized and unpersuasive. Does the author reach substantial conclusions? No. Is the description of the data used, the methods used, the experiments and calculations made, and the results obtained sufficiently complete and accurate to allow their reproduction by fellow scientists (traceability of results)? No. The authors didn't well describe the principle of variable selection. They also didn't explain their working definitions for data preparation. For example, how they define the convexity of concavity of a slope? How they quantify the human activities? How they apply Eq. 1 to normalize slope aspect (degree)? The number of slope unit is only reported in Abstract and Conclusions, which is not very logical (why not list it in Section 3.2?). Moreover, 216 slope units become 216 landslides in unit of slope in Section 4.2: the description in page 363, line 19-20 can be totally incorrect. Many gaps between methods and results make the study inconsistent, and very difficult to be reproduced by other scientists. Does the title clearly and unambiguously reflect the contents of the paper? Yes. Does the abstract provide a concise, complete and unambiguous summary of the work done and the results obtained? Yes. Are the title and the abstract pertinent, and easy to understand to a wide and diversified audience? Yes. Are mathematical formulae, symbols, abbreviations and units correctly defined and used? If the



NHESSD
1, C91–C96, 2013





formulae, symbols or abbreviations are numerous, are there tables or appendixes listing them? Please explain Ir(k+1) and Ir(k) in Eq. 2 (p. 363, line 10). Is the size, quality and readability of each figure adequate to the type and quantity of data presented? No. Please see technique correction (in the end of the report). Does the author give proper credit to previous and/or related work, and does he/she indicate clearly his/her own contribution? Not clear. The application of ANNs to landslide susceptibility analysis is not new, and the authors list about 7 citations for this topic (page 356, line 11-13) without specific comments. They also didn't highlight their own contribution. The scope of the study is also not clear. Are the number and quality of the references appropriate? Yes. Are the references accessible by fellow scientists? Yes. Is the overall presentation well structured, clear and easy to understand by a wide and general audience? No. The description for the Study area (Section 2) is limited. The authors even didn't give the size of study area. (I suggest some parts in the section 3.1 in page 359, line 4-15 can be moved to Section 2). Descriptions for distance to river and human activities are obviously missing in Section 3. In addition, the discussions are also missing the Section 5. Is the length of the paper adequate, too long or too short? The length of the paper is adequate. But I suggest the authors to reduce the number of figure by using tables or texts (Fig. 15-Fig. 18, see technique corrections). Is there any part of the paper (title, abstract, main text, formulae, symbols, figures and their captions, tables, list of references, appendixes) that needs to be clarified, reduced, added, combined, or eliminated? Yes, and please go to technique corrections. Is the technical language precise and understandable by fellow scientists? Some terms are not clearly explianed: Page 355, line 11: "landslides cannot be predicted accurately"- in terms of what? Page 356, line 22: what are "meaningful outcomes"? Page 358, line 11: what is "problem soil"? Page 359, line 24: "major semi-circular failure plane"? Page 365, line 21: "Changshougou valley" or "Changshou Valley"? Page 365, line 23: "mechanism analysis"? Is the English language of good quality, fluent, simple and easy to read and understand by a wide and diversified audience? The writing quality is not consistent (for example: Introduction vs. Data preparation). Is the amount and quality of supple-



NHESSD
1, C91–C96, 2013





mentary material (if any) appropriate? No supplementary materials for the manuscript. Technique Corrections: 1. Please use the same style (italic) for Ir(k+1) and Ir (k). Also, please change “mean squared error” to “mean square error” (page 363, line 5). 2. Figure 1: please add scales and north arrows to the two maps. Please place legends: the representation of those small points in lower map is unknown. Topographic information (contour lines and streams) can be helpful. 3. Figure 2: please combine Fig. 2 and Fig.4 (slope unit) and use different colors to separate training polygons and validation polygons. Please also enlarge the legend. In the figure caption: “Changshougou valley” is different from “Changshou Valley” in main text (p358, line 2). Please correct them. 4. Figure 3: please add an index map to locate the position/extent of Fig. 3 in the study area. Please add contour lines and streamlines and land covers, such as roads and vegetation fields. What is the straight line in the map (a profile transection)? Points are boreholes, and they are not described in the main text. From the legend, it’s very difficult to read the difference between old landslides and recent landslides. Also, by comparing with Fig. 2, many small recent landslides are not presented in Fig. 2, why? 5. Figure 5-8: Please modify their legends from stretched values to classified values so that they can be consistent to Fig.15-18. Please also add unit for those values. In the Fig. 7, the classification of aspect looks weird. In the Fig. 8, the definition of the slope shape is unknown, and the authors didn’t explain it in the main text as well. I suggest authors to combine landslide polygons and slope units to reproduce these maps, and also create maps for distance to river and human activities. 6. Figure 9: Please explain all the symbols in the main text and in the figure caption. (X1-Xn? Hidden layer? Y1-Yn? 21?) 7. Figure 12: The label of y-axis “Sum-Squared Error” is different from it in figure caption “error sum of squares”. Please correct it and also give its explanation. Please describe the red curve and the blue curve. 8. Figure 13: Please give the definition of “fitness”, and also correct the typo in the y-axis “Fittness”. Please describe the red curve and the blue curve. 9. Figure 14: Please explain this plot in more detail (the definition of “epochs and errors”). Please reorganize labels for x- and y- axis. 10. Figure 15-18: Please use a table for those plots to save the space, and



NHESSD
1, C91–C96, 2013

also list background information (frequency distributions for the entire study area) in a same table (don't miss distance to rive and human activities). 11. Figure 19: Please give legend for those gray and black polygons. Also, please combine this map and slope unit map, and use different colors to separate training and validation polygons.

---

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 1, 353, 2013.