

Interactive comment on “Comparing multi-criteria methods for landslide susceptibility mapping in Chania Prefecture, Crete Island, Greece” by M. Kouli et al.

S. Sterlacchini (Referee)

simone.sterlacchini@unimib.it

Received and published: 11 March 2013

The study analyzes the comparison of two GIS-based modeling techniques for landslide susceptibility assessment: the former is an expert-based relative weighting-rating approach, known in literature as Weighted Linear Combination (WLC); the latter is a well-known statistical/probabilistic method, referred in literature to as Weights-of-Evidence (WofE). It is really appreciable the effort made by the Authors and the topic is of high value for the scientific community. Unfortunately, some topics analysed by the Authors are not clearly explained, leading to some misunderstanding. Moreover, the paper presents many weak points:

C23

1. as a general comment, the Authors compare two different GIS-based modelling techniques belonging to different categories: WLC is a knowledge-driven method where the importance of evidence is weighted by experts. It is an expert-based blind-method, although defined as "semi-quantitative" in the study. The approach includes geomorphological and heuristic models. This approach is based on the previous knowledge of the causes of landslides in the area and a weighting scheme is applied. However, this weighting scheme is quite subjective: a "blind weighting" as suggested by Gee (1992). The reliability of the heuristic approach is directly dependent on the experience of the researchers and their geomorphological related knowledge of the factors affecting landslides occurrence in the study area. On the contrary, WofE is a data-driven method; it is the log-linear form of the Bayesian model that can be applied where sufficient data are available to estimate the relative importance of evidence by statistical means. In my understanding, these methods are too different, from a conceptual point of view, to be compared. I mean: I can compare results coming from different statistical techniques belonging to bivariate and multivariate methods and this action can provide the final-users with effective information on the robustness of results and, above all, can guarantee the repetition/replication of experiments. But, this is not the case for WLC whose results can range enormously in relation to the weighting and rating scheme defined and assigned. In order to compare results coming from so different modeling techniques, the Authors are obliged to stop the WofE analysis at the calculation of the Contrasts of the different variable classes used in the analysis instead of concluding the analysis by assessing Post Probability values, as the method can "naturally" provide.

2. Concerning landslide representation strategy, are the Authors sure that the use of points is the best solution? Have the Authors applied other strategies and assessed the results? The Authors state that all landslides are represented by points. I can share this statement for some landslide types (debris flows, above all) but I need some more information to evaluate the goodness of the Authors' choice for translational and rotational landslides, for example. And, is the pixel size 20x20 m in agreement with what has been stated by the Authors ("... small scale enforced by the great extent of

C24

the study area")?

3. From a methodological point of view, the Authors have to build different models for different landslide types, characterized by different degree of activity. If the Authors do not consider this aspect in advance, they are modeling landslide susceptibility in a generic way. The predisposing/triggering factors of debris flows are, for example, very different from those concerning deep-seated landslides. In my understanding is unacceptable to model all landslide types together: the final result will be a "generic" not-discriminant model, estimating the relative importance of evidences (predisposing factors) concerning all landslide types.

4. Sensitivity analysis has not been performed in the study. It is a strategic step in order to find the best agreement between observed and calculated results in any kind of modeling technique (knowledge or data-driven methods). This is an important weak point controlling the overall consistency of the study. That is, the tune-up of the model has not been performed adequately and this aspect controls the quality of the final results.

5. It is not so clear the conceptual scheme used by the Authors to validate the models (in terms of success and prediction rates). That is: how the Authors have used the two (80-20%) landslide datasets? Success rate is performed and presented (although no sensitivity analysis has been performed: see comment nr. 4) but no information is provided concerning the prediction rate. From a conceptual point of view, the landslide dataset has to be divided into two parts (following spatial or temporal criteria) in order to use the former dataset to tune-up the model (by a sensitivity analysis) and the latter to define the predictive power of results. From my side, the assessment of the location of future landslides is as important as the availability of a model able to describe the state-of-the-nature. For this reason, landslide susceptibility analysis has to be performed by two steps: tune-up the model; define the predictive power of the model. For this reason, it appears inappropriate the term "prediction" used in some parts of the study (abstract, paragraph 6, etc.). What has been really done in this study, given that only success

C25

rate curves are presented (Fig. 9)?

6. The evidential themes provided by the Authors are described keeping in their mind the WLC analysis. In effect, many times the Authors state that the reclassification of variables has been performed in n classes from the lowest to the highest susceptibility. This is a contradiction in terms for WofE, because it is a data-driven method.

7. Concerning the classification techniques applied, the Authors have to define them clearly for each evidential theme included in the analysis. This is a strategic feature deeply controlling the final results; one of the main aims of sensitivity analysis is, among the others, to define the variability of results in relation to the classification technique applied. Throughout the study, the classification technique has been defined only one time (page 85, line 12). WofE can provide the Authors with some interesting concepts and tools to clearly identify threshold values in case of binary classification of continuous variables.

8. Very peculiar are the "obscurities . . . confusing the WofE method results", reported at page 86: I would like to make it clear that WofE is a data-driven methods. Its response is based on the application of the Bayes' theorem on the base of the input parameters (training points and evidential themes) provided by the Authors only.

Other comments:

Page74 Lines 12-13: From my side, "the quality of the recorded landslides sample" is a crucial aspect not only for the WofE method but for all methods and modeling techniques.

Lines 22-24: Landslide susceptibility is, of course, of primary importance but it is only the first step to be undertaken to face landslide hazard assessment and risk reduction strategies.

Page 75 Lines 19-20: WofE CANNOT be listed among multivariate statistical analysis. Moreover, there are other multivariate statistical approaches that have not been listed

C26

by the Authors, besides logistic regression.

Page 76 Line 4: Please, provide some more information in order to understand why WLC is a semi-quantitative method. In my understanding it is a geomorphologically-based, heuristic "blind" method.

Page 77 Lines 3-4: Please, refer to previous comment.

Page 79 Lines 4-9: Please, rewrite this sentence making it more clear.

Page 79 Lines 21-22: There is a long debate on this topic among people supporting "stabilization effects" and those declaring "destabilization effects".

Page 81: paragraph 3.7 Elevation In this paragraph many weak points can be highlighted:

a. "... there is not any general rule relating the landslide occurrence with the elevation". If so, why the Authors force the use of this evidential theme in the analysis? How can the experts formalize "not-existing" rules? Anyway, although the elevation itself does not appear to be so important in relation to landslide susceptibility, its importance could be analyzed in relation to other parameters by introducing Unique Condition Units (UCU). By crossing elevation with other parameters ("geological and morphological parameters", as stated by the Authors) a new variable can be derived and used as a new evidential theme. This is a way, suggested by Bonham-Carter and other researchers, to avoid conditional dependence problems in using WofE. For sure, the parameters used in the definition of UCU cannot be used again in the analysis as separate evidential themes.

b. the unconditional dependence requested by WofE is violated. "... correlate as many geological and morphological factors as possible with the elevation ...". Being based on Bayes' theorem, the evidential themes used in the analysis have to be conditionally independent each other.

Line 18: "... they are uninhabited ...": why this aspect is so important for landslide
C27

susceptibility assessment?

Page 83 Lines 12-13: What does it mean "accuracy assessment purposes"?

Page 83 Lines 21-22: I beg your pardon for this comment because is partially related to my scientific activity. "... and most recently to groundwater mapping": the research group led by Masetti (University of Milan) have been applying WofE since 2001, developing many conceptual models concerning groundwater.

Page 86 From line 9 on: It would be useful to compare rates assigned by experts in WLC and weights calculated by WofE. It seems to me that the level of agreement among rates and weights is low, very low. And, to tell the truth, I am confused: I am not so sure this comparison makes sense given that the two modeling techniques are conceptually very different from each other.

Page 86 From line 23 on: As stated in a previous comment, WofE is a data-driven method and, probably, the problem is not related to WofE but to the database (training points and evidential themes used in the analysis) and to the weak conceptual model that does not include any kind of sensitivity analysis and the use of UCU too.

Page 86 Lines 27-28: "... the C_{ij} values are strongly related to the quality of the landslide sample". Results from any modeling technique are strongly related to the quality of input data. In statistics is known as GIGO effect "Garbage-In produces Garbage-Out".

Page 87 Lines 22-24: "Therefore, it is obvious that the WofE method resulted in better predictions (success rate of 87.4%) compared to the WLC method (success rate of 84.7%)". Why is it so obvious? Wasn't the WofE method confused just a few lines before?

Page 88 Lines 9-10: "So both methods provided accurate susceptibility maps ...". It doesn't seem to me.

Page 106: Fig. 6: Please, improve this figure. The transparency used in the picture
C28

does not appear to be a good graphical solution. It is very difficult (or even impossible) to detect many of the landslide susceptibility classes listed in the legend.

Pages 107-108: Fig. 7: Please, improve this figure following the comment in the previous note. Moreover, it is clear the conditional dependence problem concerning the use of WofE.

Fig. 8: Nothing appears changed concerning conditional dependence problems that propagate from one map to another. This is due to the low importance given by the Authors to sensitivity analysis.

Page 109: Fig. 9: How the Authors move from Landslide "Hazard" Index to Efficiency Cumulative Area (0-100)? Moreover: at page 85, paragraph 5 lines 10-12, the Authors state that "The landslide susceptibility maps with both WLC and WofE methods were produced within a raster/grid GIS and were reclassified into low, moderate, high and very high susceptible zones, using the natural breaks classification method". The classification technique used in this graph appears to be a sort of "geometric interval" in order to obtain the first 10% of the most susceptible area, the second 10% of the most susceptible area, etc. Please, provide some more information.

Moreover, find enclosed the revised pdf file only concerning some "technical corrections": typing errors, etc. to be provided to the Authors.

With Best Regards,

Simone

Simone Sterlacchini PhD Researcher National Research Council of Italy Institute for the Dynamic of Environmental Processes Piazza della Scienza, 1 - 20126 Milano phone: +39.0264482854 fax: +39.0264482895 mail to: simone.sterlacchini@idpa.cnr.it

Please also note the supplement to this comment:

C29

<http://www.nat-hazards-earth-syst-sci-discuss.net/1/C23/2013/nhessd-1-C23-2013-supplement.pdf>

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 1, 73, 2013.