Natural Hazards
and Earth System
Sciences

Open Access

Discussions

EGU

NHESSD

Interactive
comment

Interactive comment on "Field-based landslide susceptibility assessment in a data-scarce environment: the populated areas of the Rwenzori Mountains" *by* Liesbet Jacobs et al.

M. Jurchescu (Referee)

marta_jurchescu@yahoo.com

Received and published: 6 September 2017

"General comments"

The paper brings about an interesting discussion on a subject that landslide literature could consider as being exhausted, namely on the landslide susceptibility assessment. The study successfully responds to two major objectives, as clearly stated in the Introduction section: i) providing new and important information for understanding regional-scale landslide susceptibility patterns and its conditioning factors in a landslide-threatened area of equatorial Africa (Rwenzori Mountains' Footslopes, Uganda), characterized by very limited data and hence necessitating extensive field in-

vestigations, and ii) achieving solid statistical modeling results by thoroughly discussing several methodological choices (scale, topographic data source, spatial resolution, selecting landslide type, landslide sampling, variable selection, deciding on calibration and validation data) and grounding these on previous theoretical knowledge or on running simulations and exploring their effects. The authors propose to address landslide susceptibility in the context of a data-scarce environment by: i) constructing a regional-scale model in two phases - calibration and validation using landslide inventories in representative case-studies, followed by the extrapolation to the rest of the area -, ii) deriving separate local-scale maps for the individual case-studies and iii) using the regional susceptibility map in combination with population density for a preliminary identification of landslide risk hotspots. Given the in-depth statistical analyses, the obtained maps can be considered valuable tools for directing and implementing policy actions in the region. Results are based on several model variants (varying according to the incorporated topographic data of different sources and spatial resolutions) being run for each scale (local / regional) and subsequently evaluated. Evaluations of obtained models (e.g. in terms of performance or variables' selection) as well as the identification of the optimal ones are thoroughly conducted, being based upon statistical significance tests. Additionally, differences induced by random landslide sampling are taken into account by running 20 simulations per variant. The manuscript is well structured and written and illustrations and tables are all necessary. In my opinion, the paper is worth publishing with only minor revisions.

"Specific comments"

1. One of the qualities of the paper consists in testing the effects of several methodological choices on model performances and, especially, in the thorough comparison of these model performances in order to allow for the selection of the appropriate model variants for each level. It is highlighted that tests used to assess statistical significance of differences in performance (in terms of AUC values) highly depend on the assumptions of normality and homogeneity of variance (homoscedasticity) of the examined

values. Several steps are therefore involved: i) distributions are tested for normality, ii) homoscedasticity is tested, by also using tests which are robust to large departures from normality (Fligner-Killeen test); iii) only then significance tests are run for assessing performance differences, varying in type according to their suitability to the different categories issued.

2. An important insight is brought with respect to the scale issue, a fundamental concept in many sciences, through comparisons among regional and local models. Specifically, it is shown, not only that the regional model, when applied (validated) at the local scale, is generally outperformed by the local models, but also that it lowers its own performance when interpreted in the context of the larger scale (from 0.71 to 0.65-0.70, with one exception, according to Fig. 4). I suggest testing if this lowering from regional to local level is statistically significant and recommend that these aspects be also mentioned in the respective paragraphs (section 3.3. page 9 lines 8-13 and section 4.1. page 11 lines 10-21). This would contribute to reminding us of the scale-dependency of model outcomes, i.e. that model results derived at one scale should not be transferred at another scale, but considered valid and interpreted in the specific context of the scale they were obtained at.

3. I notice the description of data and methodology used is particularly accurate to allow reproduction. This is enhanced by the appendix describing the processing of TanDEM-X images to construct the respective DEMs as well as by the explicit Data Availability expression at the end of the paper.

4. In my opinion, other previous similar methodological questions on the impacts of varying data (of different sources, resolutions and accuracies: e.g. raised by Lee et al. 2004, Catani et al. 2013, Fressard et al. 2014) or of various modeling settings (e.g. by Poli & Sterlacchini 2007, Yilmaz 2010, Hussin et al. 2016 for landslide sampling strategies; by Heckmann et al. 2014, Petschko et al. 2014 for subdivision in calibration and validation datasets) on model performances could complement the ones already mentioned in the Methodology section (e.g. section 3, page 4 lines 15-17; section 3.1.

page 4 line 32 – page 5 line 1, sections 3.2.1, 3.2.4.). Recommended would also be the use of "e.g.", since citations here cannot be exhaustive.

5. Please also refer to motivations brought by previous papers in making methodological decisions. For example, in section 3.2.1. Landslide sampling, page 6 lines 12-13, I suggest citing Van Den Eeckhaut et al. (2006) since these authors give the same motivation for selecting only the central cells of landslide depletion zones (i.e. avoiding spatial autocorrelation and the violation of the assumption of data statistic independence).

6. If and where possible, please also briefly complement the references to outcomes of similar experiments conducted on the effect of using different data sources and spatial resolutions (e.g. Catani et al. 2013, Fressard et al. 2014) when discussing the results achieved in the current paper (section 5.1).

7. Section 3.2.1., page 6 line 6-7: I would not agree that the sentence listing landslide sampling techniques also includes: "or construct a buffer zone around (portions of) the landslide to represent the conditions under which the landslide occurred (Dai and Lee, 2003; Suzen and Doyuran, 2004; Van den Eeckhaut et al., 2006; Che et al., 2012; Hussin et al., 2016)". To my knowledge, buffer zones around landslides are drawn in order to eliminate those areas from sampling absences and not presences of landslides. Please check this and reword the phrase accordingly.

8. Figure 2 nicely presents the various methodological aspects considered in the research as well as the manner in which they were evaluated. However, it is not clear what does the "(Pre)Selection of Dependent Variables" refer to. Should there not be "(Pre)selection of Independent Variables" i.e. of explanatory variables (since the discussion on the selection of the dependent variable, i.e. landslide occurrences, is made under "Landslide sampling")?

9. I think the title of figure 4 should also include the explanation that small circles represent the mapped landslides.

10. I would suggest modifying the title of section 4.2. from "Separation of landslide types" into "Separation of landslide types and controlling variables", since this section presents results both in terms of the effect of distinguishing landslide categories and in terms of the significant explanatory variables changing according to the landslide type. This would also enable a more obvious linkage between the Results and the Discussions sections.

"Technical corrections"

- Abstract, page 1 line 20: Please delete "the" from the sentence "Topographic data is extracted from different the digital elevation models (DEMs) based....";

- Section 2, page 3 line 19: I am not sure the word "inventorized" exists, I would suggest to replace it with "inventoried";

- same correction is suggested for the next line;

- Section 2, page 3 lines 20-21: I would suggest replacing the sentence "For these two study areas a detail of the landslide inventory is given in Fig. 1" with "For these two study areas, maps of the landslide inventory are given as details in Fig. 1";

- Section 2, page 3 line 25: Please delete the word "mostly" from "In Nyahuka however, mostly shallow landslides prevail";

- Section 2, page 3 line 25-26: Please reverse the order of words, i.e. "surveyed study areas" instead of "study areas surveyed";

- Section 2, page 3 line 28-29: Please insert the word "which" as follows: "In total this inventory contains 454 landslides which were used for the susceptibility modeling";

- Section 3, page 4 line 2: I would suggest using the plural, i.e. "landslide occurrences";

- Section 3, page 4 line 3: "as the dependent variable" instead of "as dependent variable";

- Section 3.2.3, page 7 line 26: Please insert the word "of" as follows: "strong arguments support the use of elevation as an explanatory variable";

- Section 5.3, page 14 line 22: I suggest using the plural, i.e. "their effects";

- Section 6, page 16 lines 13-14: Please correct the sentence as follows: "In parallel to smaller scale landslide susceptibility studies, adequate attention should therefore also be given to investigating landslide susceptibility on the local and regional levels."

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-259, 2017.