

**Review of the manuscript “Shallow landslide prediction and analysis with risk assessment using a spatial model in the coastal region in the state of São Paulo, Brazil” by P. I. M. Camarinha, V. Canavesi, and R. C. S. Alvalá submitted to Natural Hazards in Earth System Sciences.**

Spatial data of good quality needed to evaluate landslide risk is scarce. This work presents an innovative approach using available data to assess landslide risk in an economically important and highly populated area of Brazil that has been affected by mass movements in the past. The proposed method could be transferred to other study areas. While the manuscript addresses a relevant scientific question within the scope of NHESS, it lacks a clear structure and needs clarification of some of the main aspects. With the current text the line of argumentation is difficult to follow, and therefore the conclusion is arguable. I therefore recommend publication after major revisions.

**General comments**

1. Motivation:

Clearly state what is novel in this paper. Also highlight the differences of your newly created map to the risk sector classification of CPRM, otherwise it may seem that your work in this area is redundant. What are the limitations of the CPRM risk sector classification?

2. Structure:

Especially the methodology, results and discussion section need restructuring as it is difficult for the reader to follow. Please thoroughly separate what belongs to the methodology, results and discussion sections. A good and consistent use of headings, subheadings and paragraphs will also help the reader, as a definition of the nomenclature would (e.g. susceptibility classes, themes, typologies, risk sectors, risk levels).

3. Length:

Some parts of the paper need to be shortened, especially through avoiding repetition. In some cases entire sentences are repeated (e.g. page 12), in other cases facts are stated several times. Please also delete redundant information (highlighted in pdf).

4. Input data:

The data used is limited as it is based on “available” downloads. Especially for validation purposes, data at a higher resolution than those publicly accessible through the internet would be beneficial, e.g. to assess resolution effects and validation. Moreover, why were satellite images not used to map landslide scars?

5. Weighting factors:

The assignment of weighting factors for land use and soil classes need justification. What is the degree of subjectivity in choosing these factors?

6. Discussion on resolution:

How does the resolution of your different data sets affect your results? What if the input data had 100% better resolution – how would this change your susceptibility maps?

## Specific comments

### 1. Introduction

- The introduction needs to be shortened and restructured. Some parts of the introduction are not directly relevant to the study and may be deleted, for example Page 3, lines 2-17 and 26-29, Page 4, lines 2-5, 9-15.
- Page 3, Line 23-25: You say here that a map of slope instability should provide information about spatial distribution, type, volume, speed and distance achieved by landslides. That's correct, but that's not what you deliver in this study!
- Page 4, Line 1-2: you state that the input data include triggering factors and historic landslide occurrences. I do not see where.
- Page 4 Line 2: Elements at risk – what do you mean by that?
- Page 4 line 18-20: Sentence is difficult to follow, please reword.
- Page 5, line 9: How far back reaches the historical landslide record in the study area?
- Page 5, line 14: HDI-M value *too* high – do you mean *very* high?
- Page 5, line 18: Delete first part of the sentence
- Page 5, line 20: cannot follow sentence, please reword the last part of the sentence
- Page 5, Lines 21-end: It may be interesting to pick up on these events in the discussion, i.e. did these landslides occur in areas that were classified to be of high risk according to your method?
- Page 6, lines 11-12: Are these illegal occupations captured in your approach? If not this information seems redundant.

### 2. Methodology

This section needs a consistent structure, which will make it easier for the reader. Please also make proper use of paragraphs. First, give a short outline of the section. When describing the environmental variables you may want to stick to the same order throughout, e.g. explain the particular variable, the type and source of the data you use to characterize it, and finally how the weighting factors are assigned. It would also be helpful if you could provide a short definition of the terms “class”, “typology”, “susceptibility class”, “theme”, “risk sector”, “risk level”.

- Page 8, line7: add reference
- Weighting factors for the variables “Soil class” and “Land use” are neither sufficiently explained nor supported by references in the text. It's not clear where the weighting factors come from.
- Equation 1: Please explain all terms and variables in the text
- Why didn't you analyse the satellite images for landslide scars for the validation?
- Page 9 Line 16: repetition from previous sentence
- What data did CPRM use for the risk sector mapping?
- Page 10 Line 5: risk sectors are limited to places...
- Page 10 line 20: delete 'with'
- Page 11 line 10: *the* study area / *the* literature
- Page 11 lines 15-22: It is not clear why it is necessary and beneficial to define “risk levels” in addition to the susceptibility map. It is also not clear how these risk levels were determined. Did you do it in this study? If yes, how did you do it? This classification

seems rather subjective and you do not give clear classification criteria. Please provide more detail here.

### 3. Results

I would strongly recommend separating the results and the discussion section. Please also revise the results section as it contains numerous repetitions.

- Page 12 lines 6-27: revise (whole sentences repeated!)
- Page 14 lines 4-9: While the text above is clearly part of results, this paragraph should go in the discussion section.

### 4. Discussion

I am missing a paragraph about the limitations of your method, and a critical discussion of the assumptions.

“Comparative discussion”: I do not trust such a comparison. Not only are these studies from very different terrains, but also are they based on landslide scar validation, which makes it difficult to compare to this study.

The origin of the “risk levels” is not clear. Therefore it is difficult to review and comment on your section 4.3. If you assign risk levels based on your six data sets of input data, it would not make sense to compare susceptibility classes to risk levels – as both susceptibility classes and risk levels are based on the same data and it is obvious that they agree well.

- Page 14 line 15: repetition from page 9 lines 10-18
- Page 14 line 23 – page 15 line 10: you cite a PhD thesis and a research article that are both written in Portuguese, which makes it almost impossible for the general reader to follow up these references. Could you exchange these for references in English?
- Page 15 lines 16-24: This belongs to the results section

### 5. Conclusion

- Please give a brief summary of the individual steps of your proposed methodology.
- Line 23 page 17: Please specify and discuss these considerations and assumptions, and how they may affect the proposed methodology in an additional paragraph in the discussion section (also see above). Delete “which are constantly used” as this is misleading.
- Page 17 line 24: reword

### 6. Figures and tables

- Fig 2: not directly relevant, delete
- Fig 5: this is not your work, please reconsider if this figure is necessary

- Fig. 6: if possible, please also provide weighting factors
- Fig. 9: add x axis tick labels on lowermost bar chart at the right