Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2020-134-AC2, 2020


Natural Hazards and Earth System Sciences
Discussions



# *Interactive comment on* "INSPIRE standards as framework for artificial intelligence applications: a landslides example" *by* Gioachino Roberti et al.

**Gioachino Roberti et al.**

groberti@minervaintelligence.com

Received and published: 8 July 2020

1 Introduction Reading this manuscript was interesting and stimulating. The article deals with the problem of landslide susceptibility mapping by combining different aspects, ranging from (i) the definition of a landslides classification compliant with the INSPIRE Directive, (ii) the definition of a multi-hierarchical model for the same classification, (iii) the definition of an INSPIRE scheme for landslide susceptibility and (iv) the definition of an expert-based method for the generation of maps of susceptibility to specific types of landslides in areas where inventories are scarce. In this sense the paper is adequate to the journal also because it presents innovative concepts which are adequate to international standards. Despite this I think that interventions are needed to improve the paper and make it suitable to be published.



Printer-friendly version

Discussion paper

GR: Dear reviewer, thank you for your careful revision and your insightful comments. We have addressed point by point your comments and updated the text accordingly.

2 General comments

The manuscript makes extensive use of the term artificial intelligence (also in the title). However, the type of approach used is not the classical application of artificial intelligence expected from the literature. I think it is important to clarify why this terminology is used (using the literature) or alternatively to modify the title and the text focusing more on ontology and taxonomy. In the latter case some of the titles used for the different sections of the manuscript should be changed.

GR: As stated in line 34, “Artificial intelligence” is "the synthesis and analysis of computational agents that act intelligently” (Poole and Mackworth, 2017). This definition encompass a broad range of methods and algorithms including, but not limited to, machine learning. I understand that commonly the term artificial intelligence is used as synonym for machine learning, which in turn is a term used to talk about various statistical methods. We rephrased the paragraph.

In the Method section, I think that there is a need of an introductory sub-sections which briefly introduce the flowchart of the method, even using a figure. I suggest something like: (I) defining taxonomies, (ii) defining expert-based rules, (iii) performing matching, (iv) deriving the susceptibility map. As a consequence of the flow-chart introduction the subsequent sub-sections could use a title which is compliant with the flowchart content.

GR: Good point, thank you. We have added a flowchart figure and updated the methods adding an intro. We also have rearranged Methods and Results to remove some of the repetition and add details in the webmap final results.

Sub-section 3.1.1 describe the creation of the NH classification code list for landslides. It was used for defining the expert-based rules. I wonder if this sub-section should be moved in the method section. Moreover it could be helpful to describe how the flat

representation of your classification differs from the classification of Hungr et al. (2014)

GR: We reword methods and results to clarify this point.

The method used for building the susceptibility map is based on the definition of the rules and setting of the matching scores. I wonder if these scores can have a huge impact on the final calculation of the map. I think that a discussion about how the scores are defined and which is the effect of changing those values is needed. Looking at the matching score table it seems that unmatch means -10. Is this something that should be added in the text?

GR: You are right, these scores determine the final calculation of the map. The scores are discussed at lines 115-125 and they are a measure of surprise that uses order of magnitude numbers to distinguish qualitative measures. To better explain and address the – 10, we added in the text at line 125: "In this study, an Exact match or an AKO exact match of a property with frequency "always" scores 10000, "usually" scores 9000, "sometimes" scores 1000, "rarely" scores "100" and "never" scores -10000; unmatched attributes are awarded -10 points. These scores are an arbitrary representation of degree of surprise that uses order of magnitude numbers to distinguish qualitative measures.

GR For more extensive review we refer Smyth and Poole 2004: always a proposition is "always" true, means that you are very surprised if it is false. All experience leads you to believe that it is always true, but you are leaving open the possibility that it isn't true usually a proposition is "usually" true means you are somwhat surprised if it is false. sometimes a proposition that is "sometimes" true, means you wouldn't be surprised if it is true or false. rarely a proposition is "rarely" true means you are some- what surprised if it is true. never a proposition is "never" true, means that you are very surprised if it is true. Note that these qualitative uncertainties are only the input values (i.e., as part of the models); on output we give a numerical score (both a raw score as well as a percent of the best match). This finite scale is not adequate to describe the level of matches.

In our applications we have used $\beta$ = 1000 as the value for being very surprised. (The only significant as the value for being somewhat surprised and $\alpha$ = 10000 feature of the values is the 10-fold ratios between them; 10 “somewhat surprised” is equal to one “very surprised”). Different expert models would lead to different susceptibility maps.

GR We want to stress that this paper is not about the matching system, but about the semantic and data framework provided by INSPIRE in which this method, and others, can provide interoperable landslide susceptibility maps.

The set of rules used for defining the expert-based model are important and should be visible. May be I’m wrong but I can’t find a table or a web address where those rules can be observed. I think that for a reader it could be very helpful to have access to those set of rules.

GR Links to the models are in appendix C. We also added them in the text in section 2.2.1.

In the conclusions you stated that in this study you present some landslide susceptibility maps. I would rather say that that you present a method for building a landslide susceptibility map using taxonomy and predictors/covariates and that this method is useful where accurate landslides inventories are not available.

GR: Yes, we agree. We rephrased the first sentences of the conclusions “This study presents an AI-based method to produce landslide susceptibility maps using an ontology and standardized taxonomies within the framework provided by the INSPIRE Natural Risk Zone theme. This method does not need an accurate landslides inventory to make predictions. We produced susceptibility maps for debris flow, slides in soil and slides in rock for the province of Veneto, Italy.”

About figures and tables: figure 5 is not clear. I would transform it to a table and I would let the user to go and see it in the website putting a link in the caption.

GR: Good point. We made the figure into two tables (4 and 5) , added the link in the

tables' captions, and updated the text

Specific comments

Row 45: I would remove the reference to jpeg or wav.

GR: We rephrased to: "commonly specified in data storage standards such as JPEG, or WAV"

Row 50: I think that the concept of entities and relationships should be defined a priori or some literature should be cited.

GR: We expanded the sentence and added a reference

Rows 52 -54: This sentence is not immediately clear. An example could help the reader.

GR: We added an example

Row 81: In my opinion the approach is based on the domain-expert reasoning, since rules are defined a priory. Then it mimics these rules.

GR: This is correct. The method applies rules that are set to follow expert reasoning. We rephrased to better explain.

At rows 104-105 the definitions of the model is too cryptic. I would add some examples to make clear to the reader that, if I understand correctly, a model is a set of rules defined a priori by the expert and based on the properties of the entities defined in the taxonomy.

GR: You are correct, we rephrased and added links to the expert models in section 2.2.1.

Row 110: semantic triple format and semantic network. Please cite a reference or define them.

GR: We rephrased

Row 111: “revered” or “referred” ?

GR: Sorry for the typo, we fixed it

Rows 135-136: I would use “channels” and “steep channels” in place of “streams”

GR: Ok, text is updated

Row 160: About the stream line vector layer. I suppose that only the segments in in the mountain valleys were used. But what about the starting point of these segments? This is relevant since debris flows can be triggered also in the upper part of the watersheds where channels may not have been delineated. You can discus this point for completeness.

GR: Stream line vector layer is from the Veneto Region geoportal. The dataset is not great: many channels, especially in the initiation zones, are not represented. But the idea was to use as much as possible data already available, rather than making new data. We want to provide a framework for interoperable landslide susceptibility mapping: many other methods, better data etc could be used to assess landslide susceptibility, but still can be delivered using this schema and code list extension

Row 167: is r.avaflow suitable also for slides in rock?

GR: We think so: r.avaflow allows to model fluid and solid fraction separately, you can play around with the parameters and recreate the runout of wide range of landside types, including slides in rock, assuming that the rock mass disintegrate and starts to behave as a flow-like landslide.

Rows 224-226: please use an example to describe what you have done to align the data to INSPIRE standards and explain why the same was not done for the other datasets (lakes, watersheds, etc).

GR: We rephrased the paragraph and added a figure showing the Hale Studio user interface.

Table 2: please explain how you have used the IFFI database in your model. I suppose it was used for helping to define the expert-based rules.

GR: Yes, it is correct. Expert-based model states that: Slides in Soil - has landslide type - slides in soil - always. As landslides are more likely to occur on slopes or valleys that have experienced landslides before. See links to the models in section 2.2.1

Table 3: how the Watersheds, Railroad and Road layers have been used in the model?

GR: Watersheds have been used to with Melton Ratio to classify catchment as debris flow, debris flood and flood prone Roads and rail roads have been used by assuming that roadcut and railroad cut affect slope stability, when compared to undisturbed slopes. See links in section 2.2.1.

Rows 231-234: what about the slope map? Was the map expressed in terms of classes of slope and not in degrees or percentage?

GR: Yes, slope classes are based on degrees. The matching systems adopted in this paper requires discretized data, which can be a benefit, as numbers (e.g. 37°) are harder than words (e.g. steep) by non-technical people to grasp

Row 245: 99,9th percentile of the slope units susceptibility values. I suppose. Please specify it.

GR: Yes, correct, we rephrased to: “99,9th percentile score (i. e. susceptibility values) of instances (slope units and stream buffer polygons) for each landslide type”

Row 275: please remove brackets

GR: That sentence was actually incorrect. We rephrased

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2020-134, 2020.