Review of the manuscript: Setting up the critical rainfall line for debris flows via support vector machines

submitted to "Natural Hazards and Earth System Sciences" by Y. F. Tsai, C. H. Chan and C. H. Chang3

The authors present a study on thresholds for debris flows. The topic is certainly suitable for the audience of Natural Hazards and Earth System Sciences. However, I think the paper needs major revision before publication. I hope that the following comments would give some indication to the authors on how improve the paper.

Broad comments

1. In my opinion references are too scarce and do not represent well the state of the art of the subject you are discussing. I would also have expected a broader introduction on rainfall threshold for debris flows.

2. The language needs some revision.

3. The procedure you propose follows a consistent logic. However, it is not clear how you want to use these thresholds in the territory.

Specific comments

5958 – 4: discard disaster of debris flow (in all text), use damage or other synonymous
5958 – 19/21: *lead to the amount* reformulate the sentence as it is not clear, also avoid adjectives like tremendous
5958 – 23/25: explain why you hypothesize the upward trend, maybe using references and moving the explanatory sentences that now are in the first lines of page 5959 upwards
5958 – 25/26: I think that in this sentence, but maybe in the whole paragraph, there is a problem with verb tenses
5959 – 5: Figure 1 is not necessary
5959 – 9/10: are considered
5959 – 11/12: I think it is good practice to be a little more illustrative in the introductions. If you introduce the concepts of genetic algorithm or SVM in the introduction of a paper citations and a sentence explaining briefly what they are must follow.
5959 – 16/17: the sentence is not clear, maybe a verb is missing. Moreover: what do you mean with blocks?
5959 – 20/21: check grammar
5959 – 24: Figure 2b streams with or without disaster in the legend have the same color
5959 – 25 and following: since the paper you cite is not easily available online and the definition of effective watershed area A_{15} as the area with slope greater of 15° is not so broadly known, in order to help your readers in the comprehension of the paper please place also a small definition
5960 – 3: Table 1 for me needs some explications and corrections. First of all unit measures are to be indicated with [] brackets. In the specific not (degrees) but [°]. Moreover, if S_s is the mean surface slope of A_{15} how comes that the values, in degrees, are lower than 15°?
5960 – 5: discard statistical
5960 – 7: discard statistical
5960 – Data processing paragraph: I think that your notation is not clear and do not help the comprehension of these equations. You use j both to differentiate vector F_i and F_j and as an alternative counter for M. It allows confusion in my opinion. Moreover, as best practice, you have to explain each equation. I know that they are pretty standard but if you write them you have to devote two lines in explaining them.
5961 – Clustering analysis paragraph: In this paragraph it should be clearly stated why you apply this procedure. The algorithm separates groups of similar processes but it seems to me that this is independent from the location of the stream so it seems that your output would provide thresholds for each stream not on regional basis, which could be precise and useful but should be explicitly stated.

5963 – 8/12: I would advise you to discard Fig.4 and Fig.5a since they represent a well-known problem in scientific literature, use references. In order to explain better SVM and why you pass in a higher dimensional space, you can modify figure 5b using data as fields and explicit better in the texts that this helps you with overlapping.
5963 – 14/15: A *that* is missing
5963 – 20/21: A verb is missing
5964 – 2: in Figure 7 the blue dots are very difficult to detect, try to switch to yellow or other brighter colors
5964 – 9: I would not use *likely*
5964 – 18: *for* debris flow (in all text)