

Dear Reviewer,

Thank you very much for your kind comments on our paper, I will reply your comments one by one, as follows:

**Question [1]:** .....however, for better clarity it could be useful to highlight the practical/operational advantages and disadvantages of the geophysical approach over current early warning system based in contributing factors, besides forecasting results.

**Reply:** The authors agree that it is better to describe the operational advantages and disadvantages of our new system. The significant advantages are believed to be the DF watershed that is used as the basic forecasting unit in our system and are listed between the lines 20-25 in Page 17 in the discussion paper. Of course, the disadvantage is that the calculation of initial soil water content value for DF forecasting is time intensive. This situation will be not beneficial to operation. We added a sentence in the 29<sup>th</sup> line in Page 17 in the Discussion Paper in order to address this disadvantage as follows: “Despite the fact that the new EWS has higher precision, it still should be noted that the calculation of initial soil water content for DF forecasting in our new system is time intensive. This situation is not beneficial to the operation and this is the disadvantage comparing to the current early warning system based in contributing factors.”

**Question [2]:** A revision of the English language is required in order to improve readability.

**Reply:** To improve the readability of the manuscript, we have sent it to an English native speaker to polish the English language.

**Question [3]:** For better readability of figures 10, 12, 14 and 15 it should be clearly stated in the text (lines 5-12 pp 5863) how warning levels were homologated. Does current EW system based in contributing factors uses the same warning levels as the proposed one based on water-soil mixture density?

**Reply:** The authors have stated the Figure 10 in line 7 in Page 16, Figure 12 in line 11 in Page 17, Figure 14 in line 21 in Page 17, and Figure 15 in line 21 in Page 17. In order to more clearly state Figures 14 and 15, we added a sentence in line 21 in page 17 as follows: “Figures 13 and 14 are the transforming results and they indicate that...” In this comment, another issue proposed by the Reviewer is that whether current EW system based in contributing factors uses the same warning levels as the proposed one based on water-soil mixture density. Exactly speaking, the warning levels are not completely same due to different algorithms adopted in the two forecasting models. However, the warning levels used in the two models represent a tendency that the debris flow formation probability is higher and higher as the color ranging from blue to red. So despite the warning levels are not completely same, the warning results should be not big differences when a grid or DF watershed is marked the same color by the two models.

**Question [4]:** Table 1 should clarify whether it applies only to the new forecast method or both methods. Table 5 and related text (lines 25 to 29 pp 5863) should use standard terminology of true positives/false positives true negatives/false negatives. The authors may be interested in presenting these results in form of confusion tables (Fawcett, Tom (2006). "An Introduction to ROC Analysis". Pattern Recognition Letters 27 (8): 861– 874. doi:10.1016/j.patrec.2005.10.010). The indexes derived from such tables may improve the clarity of their points.

**Reply:** Table 1 is specific for our new forecasting model and corresponding system. In order to clearly state this situation, we added a sentence in line 9 in Page 9 as follows: “This Table is specific for our new model, and which is not suitable for any other DF forecasting model mentioned in this study.” As for the standard terminology for expressing the forecasting effectiveness, the

authors would like to thank the reviewer for providing the reference paper. After carefully reading this reference paper, we decided to use the standard terminology of true positives/false positives true negatives/false negatives instead of the expression of “Prediction failure” and “False prediction”. We used “false positive rate” to replace “false prediction rate”, and used “false negative rate” to replace “failure prediction rate”.