Dear Editor

The paper has been furtherly modified to include most of suggestions provided by reviewers, as hereafter indicated:

Reviewer #1

Although the manuscript has been improved enough to be accepted, english language could be refined and I have included in the attached file some suggestion, that may make it more clear and attractive to readers.

We thank the Reviewer for his/her advices. We have accepted most of the suggested changes. Here follow the answers to the major issues raised.

(Page 3, lines 29-31) *"cause variable" and "effect variable" are correct terms, but "independent variable" and "dependent variable" are more esplicative and more common terms.*

We know that in Mathematics the terms "independent" and "dependent" are more formally used, but here we want to stress the cause-effect relationship, so we prefer "cause" and "effect".

(Page 6, from line 22 onward) *Too many examples: I think that readers, by using a single example, are able to generalize this simple concept. The text could be drastically shortened.*

The other Reviewer suggests to even support such a description with some examples of application, as he probably thinks that this concept is not so obviously generalized. So, we decided to leave this part as it was.

(Page 7, lines 6-7) *Also in fine grained soils, depending on water content and slope steepness, earthflows can take many years or just minutes to move a significant amount. Please, specify the adopted classification criterium of landslides*

We added the word "some", as we see Reviewer's point. We also specified the criterion of landslide classification we are adopting.

(Page 12, lines 1-2) *Please, minimum and maximum sizes (thickness and area) of the landslides must be included; activation dates of the landslides must be included.*

We added Table 1 in order to indicate the site, the dates and volumes of landslides

(Page 12, lines 7-13) *The rainfall aggregation time (e.g. the hourly scale) can be more relevant than antecedent rainfall (for example, 30 mm in one hour can trigger a shallow landslide, while 60 mm in 24 hours do not trigger a shallow landslide on the same slope, despite a high soil moisture content). Please, a brief discussion must be included in the text with reference to rainfall aggregation time.*

We see Reviewer's point, but it pertains in any case to the triggering rainfall, and not to the initial slope moisture conditions, which instead depend on antecedent precipitations. Both 30mm in 1 hour and 60mm in 24 hours produce different effects on the same slope after different antecedent precipitations. Furthermore, the aggregation time for the rainfall is practically dictated by the available data set (often only daily rainfall is available).

(Page 13, lines 34-35) *why 0,4, 0.6 and 0.8? How many attempts are need to define these thresholds (I suppose based on the ratio false/missing alarms)?*

We have added the reference where the choice of the thresholds is described. We have also rephrased the part where we describe the choice of the sensitivity of the EWS setting, to point out that ir is also related to the choice of the thresholds.

(Page 15, line 18) *Please, a brief discussion must be included in the text with reference to rainfall aggregation time (also with reference to the previous example of the Sarno landslides).*

The aggregation time for the rainfall is practically dictated by the available data set. Often only daily rainfall is available, in this case hourly data were available.

(Page 15, line 35) *why 0.75 and 1.00?*

It is an example choice, and we have now specified it. We have also rephrased the part where we describe the choice of the sensitivity of the EWS setting, to point out that ir is also related to the choice of the thresholds.

(Page 16, lines 9-13) *The mobility function (and the threshold Ycr) is the same for the two cases (22-25 sept 1993 and 12-15 oct 2000)? In my opinion the 2000 case must benefit from the knowledge of the condition of 1993 landslide activation.*

We used the same threshold for the two cases for the sake of simplicity, and because we wanted to show how a model previously calibrated can assure good performance when operated in real time. Of course, during the actual operation of an EWS, the predictive model benefits of continuous calibration update by incorporating the new information collected every time a new landslide event occurs.

(Page 17, lines 6-7) *How many kilometres?*

The farness of the meteorological station was indicated in the text

(Page 17, line 7) *Can a single landslide activation be used to calibrate and validate the model? A different combination of factors causing the slope instability can produce the same result. I suggest to discuss this consideration, also including idrogeological and geotechnical data. In addition, data must be provided in order to ensure repeatability of the experiment.*

The physical based approaches do not require calibration based on case-studies, as their parameters, having a physical meaning, may be quantified by means of laboratory tests on specimens, so as it was carried out for the present case. In practical applications, they should be adopted right when the number of the well-documented case studies is so poor to make unadvisable following statistic approaches. In this work the landslide case study, unique for nearness of the meteorological station to the landslide area and availability of hourly resolution records of rainfalls, was adopted to validate an already calibrated model. Validation refers not only to the prediction of landslide occurrence at the specific landslide time but also to prediction of landslide lack over a meteorological window 10 years long. Hydrological and geotechnical data are reported in the indicated reference (Pagano et al., 2010).

Reviewer #2

The Authors followed some of the comments and suggestions provided. The revised manuscript results more clear and the introduction has been improved. However few issues are still unsolved from my point of view.

We thank the Reviewer for his/her further comments. Here follow the point by point answers to the raised issues.

1. Predictive model and/or interpretative model?? Following the authors' response: "The predictive tool is defined in the Introduction, at page 3, lines 27-29, as the "interpretative model". However the sentence of pag. 8 lines 26-27 ("The second step of the development of the predictive tool is choosing the interpretative model.") is meaningless and should be revised. I agree with the authors' response:" the misunderstanding is due to the use of different names (e.g. "interpretative model", "predictive model", predictive tool") throughout the paper. This problem should be fixed, because from the sentence at pag. 8 lines 26-27, seems that predictive tool is different from the interpretative model. Basically, I would suggest to change, in the text and in the title, the words "predictive tool" with the more pertinent "predictive model" or "interpretative model". The word tool generally means "something (such as an instrument or apparatus) used in performing an operation" (source: Merriam-webster dictionary).

We followed the suggestion of the Reviewer with regards to the misleading use of the attributes "interpretative" and "predictive", which refer to the same model, so we now use always "predictive model". About the use of the word "tool", according to the proposed dictionary definition, we use it all the times we refer to the operational instrument which makes use of the model predictions to activate risk mitigation measures. We checked the manuscript carefully, and left the word "tool" only where it is used properly.

2. I am still missing how and why the predictive tool can be generalized to any water-related early warning system. From my point of view the generalization of a concept is a bit more complicated and request an important bibliography analysis in order to support the generalization (which is still the aim of the authors) with different case studies. For instance, how can be scientifically demonstrable that "the fame proposed throughout the paper may be referred to all scales and other types of natural hazards" (pag.4 lines 10-15; lines 22-24), if any example, in this regard, is provided? To strengthen the thesis, the authors could have described different real examples and/or cases from literature. The three examples provided all refer to shallow landslides at a local scale.

Indeed we already provide examples of how the proposed frame is applicable to earthquakes (page 6, lines 22-33), snow avalanches (page 7, lines 8-14) and to water-related hazards characterized by different scales (landslides: from page 6, line 34, to page 7, line 7; floods: page 7, lines 15-23). Furthermore, the other Reviewer even suggests to shorten/eliminate such descriptions, as he thinks "that readers, by using a single example, are able to generalize this simple concept". So, we decided to leave this part as it was.

3. Conclusions are still a bit general. They could be improved describing which is the lesson learnt.

We have added a new paragraph at the end of the Conclusions section, in order to address the issue raised by the Reviewer.

Specific comments:

pag. 2 lines 17-19: "significant examples of operational early warning systems are currently found in the field of floods, landslides, snow avalanches, earth fill failures." Please provide a reference for each one to demonstrate the existence of such systems.

Several references are already given few lines below, where the features of such EWS are discussed case by case: for floods, Rabuffetti and Barbero (2005), Cranston and Tavendale (2012), Alfieri and Thielen (2015), de Saint-Aubin et al. (2016); for landslides, Keefer et al. (1987), Ponziani et al. (2012), Ortigao and Justi, (2004), Chleborad et al. (2008), Baum and Godt, (2010), Capparelli and Tiranti (2010), Tiranti and Rabuffetti (2010), Segoni et al. (2014), Tiranti et al. (2014), Piciullo et al. (2016), Eichenberger et al. (2013), Pumo et al. (2016); for avalanches, Bakkeoi (1987), Liu et al. (2009); for earth fills, Pagano and Sica (2013), Ma and Chi (2016). In addition, the review by Alfieri et al. (2012) is suggested for EWS dealing with water-related hazards.

pag. 3 lines 32-34: "…assessing system safety conditions". is it really the aim of the predictive model? Explain why.

As written (page 3, lines 34-35), the predictive model links "cause and effect variables, allowing to catch the evolution stage of the phenomenon". When used as the predictive tool of an EWS, its aim is exactly to quantify if the monitored system is safe or not, and how far from being unsafe it is.