Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2017-427-RC2, 2018

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Probabilistic landslide ensemble prediction systems: Lessons to be learned from hydrology" *by* Ekrem Canli et al.

Anonymous Referee #2

Received and published: 13 January 2018

General comments

Landslide early warning is a very interesting and relevant topic. However, in my view the authors fail to deliver what they promise. After reading a 10 page long introduction, I was convinced that the authors were going to introduce a new framework/method to incorporate ensembles of weather forecasting for landslide prediction, similarly to what has been done for hydrological forecasting. However, that is not the case. The authors selected a landslide physically based model, which they run considering uncertainty in some pre-selected model parameters. Nowhere, did the authors considered uncertainty introduced by the weather forecast that is essential in early warning. While it is valid to use a Monte Carlo approach to evaluate uncertainty introduced by the parameters of the landslide model (many studies have done this, both in the hydrology

and the landslide community), this is different from assessing uncertainty in a weather forecast (by using an ensemble for weather forecasting, provided by operational forecasting system such as those listed on page 6, lines 12-17) particularly relevant for early warning systems. So my main concern with this paper is that while the reader is given the impression that the paper is about uncertainty in weather forecasts to improve landslide early warning systems, the paper is about something else, i.e. uncertainty in landslide model parameterisation. Why then writing so much about early warning and ensembles for weather forecasting (one example is the entire section 4, but I will provide more examples in the Specific Comments section) if the paper is about something else? I believe that this study could be a really interesting if the authors have indeed used landslide physically based models together with many different initial conditions of rainfall (based on a possible weather forecast), to enhance a landslide early warning system at a regional scale. Given the results presented, in my opinion this paper needs to be restructured so that the first six sections reflect what the paper is really about, that is uncertainty related to the geotechnical parameterisation of the landslide model, as stated on page 10, line 5.

Other major comments include:

- Clearly state what is novel about this paper and its direct relevance for early warning. Monte Carlo simulation as a way to incorporate parameter uncertainty is not new (just few examples in a landslide context, among many, Haneberg, 2004; Cho, 2007; Melchiorre and Frattini, 2012).

- Lack of justification on why only certain model parameters were considered uncertain and not others.

- Not enough detail is provided on the probability distributions used for the uncertain parameters, which is essential for the reproduction of the results.

- The paper lacks a proper structure. It seems more like a thesis chapter, than a paper for a journal. The Introduction is too long (six separate sections), the discussion is incomplete, and the main conclusions that arise from this study are non-existent. What the authors call Conclusions section, reads more like a Discussion section. The fact that the study lacks a Method section, jumping straight from Introduction to Case Study, makes the paper more difficult to read.

Specific comments

Page 1, line 13: What do the authors mean by "larger scales"?

Page 1, line 14: Why convective-scale numerical weather predictions specifically and not other numerical weather predictions? Please also briefly explain what convective-scale numerical predictions are.

Page 1, lines 17-20: Is this study going to answer any of the future research directions identified by the authors in the Abstract? Or are the future research directions identified here, gaps that need to be addressed in the future? If the first, please make clear in the Abstract what the main conclusions of this study are and how this study contributes to the points raised. If the latter, before pointing the future directions of research, please make clear what the main conclusions and contributions of this study are.

Page 1, line 31 – Page 2, lines 1-2: The authors say that "For natural hazard types with a rapid onset (. . .) rainfall can be considered as the main triggering mechanism." What about earthquake triggered landslides? They also have a rapid onset and are not triggered by rainfall. Or wildfires, etc.?

Page 2, lines 7-12: The authors say on page 2 that there are multiple reasons why numerical weather predictions are not used within the landslide early warning community. One of those reasons pointed by the authors is the "the complexity of single landslide detachments: the same landslide triggering event does not necessarily cause other landslides as the time between propagation stage and the collapse phase may vary significantly based on differences in local conditions (. . .) and spans from minutes (. . .) to years (. . .)". While it is true that the same rainfall event may result in a landslide

in certain places (and for certain initial conditions) but not in others, and that the collapse phase may vary from minutes to years, I do not fully understand how this relates to the lack of uptake of numerical weather predictions by the landslide early warning community. Please make the link clearer to the reader.

Page 2, lines 25-29: The authors claim that "This paper reviews and summarizes concepts of ensemble prediction systems (EPS) in hydrology and how those can be translated to be applicable also in process-based landslide early warning systems. A strong emphasis is put on how to deal with spatial uncertainties by demonstrating the benefits of probabilistic model application which does not eliminate uncertainty, but it explicitly introduces in into the model results." While the authors provide a case study where they consider spatial uncertainties and how these uncertainties impact the model output (Factor of Safety), I do not agree that the authors show how ensemble predictions systems can be used in process-based landslide early warning systems. For early warning systems, it is essential to consider the uncertainty in weather forecasts, and not just the uncertainties introduced by model input parameters as the authors do in their case study. Please make the text here more accurate, so that it reflects what is later delivered.

Page 2, lines 31-33, Page 3, lines 1-2: The authors state at the end of Section 1, that the aims of the paper are: "a) to critically evaluate the current state of physically based landslide early warning, its limitations and possible ties to hydrological forecasting; b) on this basis, to foster cooperation across disciplinary boundaries to bring together scientists from different fields to pursue research based on forecasting experiences gained in the last couple of years". Is this paper a literature review on this topic, or does it aim to introduce something new? If the first, please make it clear. If the latter, please clarify what are the research questions that the authors aim to answer.

Page 3, lines 29-30: Streamflow measurements for extremely high conditions are difficult to obtain and far from accurate. Please consider rephrasing the sentence in your manuscript.

Page 4, lines 10-11: How does this sentence link to what comes before? Why do the authors jump to data collection?

Page 4, lines 12-25: Why do the authors talk about underestimation of landslide losses here? Why is this relevant for the overall argument that the authors making? Please clarify.

Page 4, lines 27-28: Perturbations to model parameters is one thing, perturbation to starting/initial conditions is another (as it is clear from the definition given by WMO that the authors quote in page 3, lines 5-11). It is crucial to clarify what this study is about. Up to this point, the reader is made to believe that perturbed initial conditions (reflecting uncertainty in weather forecasts) are going to be used (and maybe also different combinations of physical parameterisation schemes of the landslide model). However, that is not the case. The study is about uncertainty in the physical parameterisation of the landslide model. Please note, that I acknowledge that the authors make a more clear distinction later on page 5, lines 5-9, when they define the term "ensemble prediction" for multi-parameter and multi-model predictions. But my point is that, the reader does not know what the paper is about. And if the paper is going to be about parameter uncertainty (not clear at this point), why do the authors spend so much time talking about forecasting, weather ensemble forecasting and early warning?

Page 5, line 10: I do not agree that there have been only a few attempts to use ensemble techniques in landslide research (assuming here that the authors are talking about multi-parameter and multi-model ensembles). Some examples include: Haneberg, 2004; Rubio et al, 2004; Cho, 2007; Melchiorre and Frattini, 2012, Arnone et al, 2014. Some of the references provided by the authors later (page 5, line 27) are possibly also good examples.

Page 5, line 12: Once again when the authors say that "None of them, however, incorporate ensemble techniques in real-time application", makes me believe that this study will address that, what in my opinion is not the case.

Page 5, lines 28-32: Similarly to the previous comment, when the authors say that "Haneberg (2004), Park et al. (2013), Raia et al. (2014), Lee and Park (2016) and Zhang et al.(2016) treat soil properties at regional scale applications in a probabilistic way by randomly selecting variables from a given probability density function, mostly by means of Monte Carlo (MC) simulation (. . .) None of those probabilistic approaches are operated in spatial real-time-early warning systems, not even on a prototype basis", makes me believe that this study will go beyond using Monte Carlo simulations of soil properties randomly sampled, and provide a method/case study/framework to be used it in real time warning systems. That does not turn out to be the case.

Page 6, line 28: What do the authors mean by hydrological applications?

Page 7, lines 8-26: Again, why do the authors spend so much time talking about quantitative precipitation forecasts (QPF) from numerical weather predictions, including giving the example of flash floods that require QPFs with 1-6 hour lead times, if the study does not use such QPFs?

Page 8, lines 5-6: This is a confusing sentence mixing equifinality and nonlinearity. Please consider rephrasing it.

Page 8, line 31: The concept Factor of safety is introduced here (page 8, line 31), and defined later on page 11, lines 21-24.

Page 9, lines 11-13: The sentence "Commonly, calibration will improve the reliability of forecasts (i.e. the match of the target variable or forecast probabilities to frequency of observations of the event) but reduce the resolution of the forecast (the ability to discriminate whether an event will occur or not)." appears twice in this manuscript (page 9, lines 11-13, and page 18, lines 17-19). Please do not repeat the exact same text twice in the same manuscript. It is also unclear what the authors exactly mean. I have not seen before "resolution of forecast" being defined as "the ability to discriminate whether an event will occur or not"

Page 9, lines 13-14: “calibration will improve forecasts of common events, but will also lead to the underprediction of more extreme events.”- Please add reference(s).

Page 9, lines 17-26: This paragraph is confusing. At the beginning of the paragraph the reader is given the impression that the authors are going to talk about validation. But then the paragraph seems to be about calibration.

Page 9, lines 24-25: What do the authors mean by “dependence on the model structure”?

Page 9, line 33 and page 10, lines 1-2: What is the point of providing these two references here, without a brief description of the measures? If those measures are relevant for the case study, please consider giving a brief description here. If not, I am not sure why the authors introduce the references here.

Page 10, line 3: After 10 pages of introduction (not sure why it has to be so long!), it would be helpful to state clearly what are the challenges that are going to be addressed in this paper and the research questions, before moving to the case study section. So far, this reads more like a book/thesis than a journal paper. Jumping from introduction to case study is also not very common.

Page 10, line 4: What do the authors mean by “simplified ensemble modelling”? What has been simplified?

Page 10, lines 5-7: It is unclear what is done in this study regarding point c) “how infrastructure data can further supplement early warning procedures in an exposure context.” Later on in the study (page 14, lines 3-5), the authors overlap the results of the landslide model for a past rainfall event with Open Street Map. However, given that this is based on a past rainfall event (and not forecasting weather) how can this be used to supplement early warning procedures? Please rephrase point c) to reflect the results shown later on, or consider deleting point c) from this list.

In here the authors state more clearly what the aims of this case study are. Taking into

account that uncertainty in weather forecasts in landslide prediction is not part of those aims, why spending so much text up to here on that that topic?

Page 10, lines 26-27: Please add references after “physically based models can be quite commonly found to evaluate rainfall-induced landslide susceptibility at the regional scale”

Page 12, lines 4-7: The description of the model setup is too generic. For example, what is the time step used? How many parameters does the model have?

Page 12, lines 4-7: What are the properties of the normal distribution that the authors sample from (i.e. mean and standard deviation) and how were they derived? On what grounds did the authors select a normal distribution and not any other distribution?

Page 12, lines 7: Why 25 model runs? Such a small number may not adequately represent the parameter space.

Page 12, line 8: What is the “initial model run”? Is it the first of the 25 model runs?

Page 12, lines 9-12: Please consider rephrasing this sentence, as it is confusing. The authors say that the probability of failure of a given cell is equal to dividing the number of unstable raster cells by the number of model runs. It is confusing as for "this raster cell" they count the "number of unstable raster cells". I suppose that the authors mean the number of simulations that lead to FoS<1 for that specific raster cell, divided by the total number of simulations (i.e. 25). Is that the case? If so, please change the text to reflect that.

Page 12, line 21: Is the data from Tofani et al. (2017) mentioned here used by the authors to derive the parameters of the probability distributions of the uncertain parameters (i.e. mean and standard deviation of the normal distributions)? If so, and as mentioned in a previous comment, please provide the values of mean and standard deviation of the distributions and how those values were determined.

Page 12, lines 25-26: “(. . .) the boxplots suggested normal to lognormal parameter distributions. This is a common observation and might be a result of the central limit theorem". A lognormal distribution is not the result of the central limit theorem.

Page 12, line 25: If the boxplots suggest normal to lognormal distributions, why did the authors decide on normal distributions? Please justify the choice made. Is the observation "the boxplots suggested normal to lognormal parameter distributions" valid for all model parameters? Please detail which parameters show a normal distribution and which parameters show a lognormal distribution.

Page 12, line 30: Why do the authors introduce GLUE here? GLUE indeed involves Monte Carlo simulation, but there is more to GLUE than that. In this study, Monte Carlo simulations are performed, but GLUE is not. So I do not see the need to mention GLUE here. It may only contribute to confuse readers that are not familiar with GLUE.

Page 12, lines 31-32: Are the parameters sampled from a predefined parameter range or are they sampled from normal distributions? Please clarify. Page 12, lines 1-2 and line 33: Please justify why only uncertainties related to soil depth, effective cohesion and effective friction angle were considered. How do the authors know what are the most influential model parameters, without having carried out a sensitivity analysis?

Page 13, line 6 (caption of figure 2): What is "state" function?

Figure 2 (page 27): As mentioned in a previous comment, please provide the parameters of the normal distributions. Some of the points in Figure 2 do not seem to come from a normal distribution (e.g. for friction angle the distance between the 25th percentile and 50th percentile is quite different from the distance between the 50th percentile and 75th percentile).

Page 12, line 7 and page 13, line 11: Is the model run for a specific rainfall event? Why 3 hourly time steps? Was 3 hours the duration of the rainfall event? Please provide enough detail on the rainfall event used to run the landslide model and why that specific event has been selected. Page 13, lines 13-14 and lines 20-21: This is where the study

falls short, i.e. show how numerical weather predictions and related uncertainty could be incorporated in a real-time application. If the landslide model is run with rainfall over the last three hours (page 13, lines 20-21), instead of a rainfall forecast, this cannot be used for early warning. But wasn't early warning the main selling point of this study?

Page 13, line 20: Are there 24 or 25 model simulations?

Figure 3 (page 28) and page 13, lines 23-25: The 24 figures are not very helpful, as the reader cannot see the differences between the maps. What is the message that figure 3 is trying to illustrate? Is it essential to show these 24 maps, given that the reader cannot see much?

Based on the figures provided I cannot see whether "the results indicate quite significant changes across individual members, but also quite high similarities although parameters change drastically between some of the members" (page 13, lines 23-25)

Page 13, line 32: Please make clear what uncertainties are accounted for, given that only certain model parameters were considered uncertain, and other uncertainties such as model structural uncertainty and rainfall uncertainty, were not considered.

Page 14, line 2: What do the authors mean by "this specific time"? Please clarify.

Page 14, lines 7-8 and figure 4 (page 29): The results shown in this study refer to a specific rainfall event. This should be made clear, namely in the main text and in the captions of figures 4 and 5. If it does not rain, the values shown in figures 4 and 5 are no longer the probability of failure. And if it rains a lot (meaning much more than the rainfall event used to run the landslide model) the probability of failure is much higher than what is shown in the figures. Please make clear what the probability of failure refers to.

Page 14, line 11: It is incorrect to say that a narrow ensemble spread is an expression of equifinality.

Page 14, lines 12-13: There is not enough evidence in the results shown in this manuscript to make such a statement, as the authors did not run the model for different rainfall events. How can the authors know that for a different rainfall forcing the location of possible slope failures is not different?

Page 14, lines 28-29: Indeed, but there are equally also many landslide initiation points that do not correspond to high failure probability. Of course, this may have to do with the rainfall input used. But that is my point - the probabilities shown in the map only have meaning for the specific rainfall event used, which the reader does not even know what it was.

Page 14, line 32: “there are still many accounted uncertainties”: Please list the unaccounted uncertainties (or at least some of them).

Page 15, line 6: Why are some real landslides missed?

Page 15, lines 14-15: What do the authors exactly mean by “spatial confidence buffer”? Are these the coloured areas in figures 4 and 5? How does the “spatial confidence buffer” show a narrow ensemble of spread? Spread of what? How does that relate to an equifinal result? How does that relate to slope angle?

Page 15, lines 14-14: How do the authors know that slope angle is the main predetermining factor based on their results? No sensitivity analysis has been performed, to make such a statement. If this statement is based on other studies, that needs to be made clear. It is important to highlight that the most influential parameter highly depends on the probability distributions used.

Page 15, lines 17-18: The statement “no matter what the geotechnical or hydraulic input parameters are, it will be always the same slope segments that will result the highest slope failure probability” is not accurate. Geotechnical and hydraulic input parameters may still matter, even if slope angle may have the greatest impact on the model results. A slope with a certain angle (and assuming the same rainfall event), may fail for a certain combination of geotechnical and hydraulic parameters and not fail

for another combination of geotechnical and hydraulic parameters.

Page 15, lines 18-19: It is unclear what the sentence “Slope failure probability will ultimately vary only based on the dynamic component (here: rainfall) or if a spatially distributed soil depth map is provided” mean. Does failure of probability only depend on rainfall and soil depth map? What about the other aspects, such as slope angle, friction angle, cohesion etc?

Page 15, lines 26-27: How does that add large errors? Please expand. How does that compare to the approach introduced in this paper, where only a pre-selected rainfall event has been used? Page 16, lines 13-14: What does the sentence “However, narrowing down uncertainties is a good first step, but not the be-all and end-all of ensemble approaches.” mean? What does it mean “be-all and end-all of ensemble approaches”? How do the authors suggest narrowing down uncertainties? Page 16, lines 23-25: But the authors do not do use physically based predictions with blends of most recent quantitative precipitation estimates either. The authors fail to show how the approach/results they present in this study can be used for early warning. For example, could the model be run fast enough to be used for early warning based on the weather predictions? This is just a very simple example.

Pages 16 and 20, Conclusions section: Line 30: The Conclusions section must start by clearly stating what was learnt from this study, i.e. what the actual conclusions of this study are. Only after this, the authors should discuss any challenges and/or future direction.

Large parts of the Conclusions section should be moved into the Discussion section.

Page 17, line 8: Please provide references after “(...) probabilistic treatment of input parameters for regional model application has seen a rise only in the last couple of years.”

Page 18, lines 2-3: “when using literature values instead”: It is unclear what the authors mean. Are the authors saying that literature values should be used instead, and that sample measures should be discarded? Please clarify.

Page 19, lines 15-17: I am not sure that this sentence makes sense. Furthermore, the results presented in this study do not allow the authors to make such a strong statement.

I would expect that if we are looking into a smaller area, the variability of certain soil properties (or more generally input parameters) is smaller than if we would be measuring the same properties across a larger area. But most importantly, the authors do not show results to back up their statement.

Page 19, line 20: I do not follow the argument. According to the authors, is computation problem still a problem nowadays or not?

Page 20, lines 5-6: People have been using HPC in landslide modelling – please check the literature.

Technical corrections

Page 1, line 16 – "how ties to. . ." – Please rephrase.

Page 2, line 19: "Another reason for the negligence of physically based forecasting initiatives (. . .)" – Please rephrase

Page 2, lines 24-25: "The hydrological community has recently adopted to those advancements (. . .)" – Please rephrase.

Page 2, line 29: "in into" – Please correct.

Page 3, line 26: In the sentence "One reason why landslide forecasting is seemingly more challenging can be (. . .)", please state what landslide forecasting is more challenging compared to, i.e. "One reason why landslide forecasting is seemingly more challenging THAN X can be (..)"

Page 6, line 13: Delete the comma after "In".

Page 7, line 16: "longer lead times (. . .)" Longer relative to what?

Page 8, line 31: "(. . .) when if" – please correct.

Page 9, lines 17-18: What do the authors mean by "either-or-situations"?

Page 14, line 30: What situation are the authors referring to? Please clarify.

Page 14, line 2: Please clarify what "This" refers to, or in other others explain what is quite detrimental. Is it explicitly accounting for uncertainty?

References

Arnone, E., Dialynas, Y. G., Noto, L. V. and Bras, R. L. (2014). Parameter uncertainty in shallow rainfall-triggered landslide modeling at basin scale: A probabilistic approach. Procedia Earth and Planetary Science, 9, 101-111, doi: 10.1016/j.proeps.2014.06.003.

Cho, S. E. (2007). Effects of spatial variability of soil properties on slope stability, Engineering Geology, 92(3-4), 97-109, doi: 10.1016/j.enggeo.2007.03.006.

Haneberg, W. C. (2004). A rational probabilistic method for spatially distributed landslide hazard assessment, Environmental & Engineering Geoscience, 10(1), 27-43, doi: 10.2113/10.1.27

Melchiorre, C. and Frattini, P. (2012). Modelling probability of rainfall induced shallow landslides in a changing climate, Otta, Central Norway, Climatic Change, 113, 413–436, doi:10.1007/s10584-011-0325-0.

Rubio, E., Hall, J. W. and Anderson, M. G. (2004). Uncertainty analysis in a slope hydrology and stability model using probabilistic and imprecise information. Computers and Geotechnics, 31(7), 529-536. doi: 10.1016/j.compgeo.2004.09.002

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427, 2017.