Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2017-213-RC1, 2017

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Development of a Precipitation-Area Curve for Warning Criteria of Short-Duration Flash Flood" *by* Deg-Hyo Bae et al.

Anonymous Referee #1

Received and published: 24 July 2017

Review of

'Development of a Precipitation-Area Curve for Warning Criteria of Short-Duration Flash Flood' submitted to NHESS by Deg-Hyo Bae, Moon-Hwan Lee, Sung-Keun Moon

Dear Authors,

In this manuscript you propose a model for predicting (and warning of) flash floods in parts of South Korea. You argue that a soil-water model is an important component for better deriving conditions that lead to flash floods. You also include the concepts of flash flood guidance (FFG), threshold runoff, and simulations of virtual rainfall to arrive at a precipitation-basin area curve that helps predict flash floods. The backbone of your approach seems to be a classifier that allows you to 'predict' flash floods from

Printer-friendly version

Discussion paper

time series. One could interpret your precipitation-area curve as a decision boundary, although you do not explicitly investigate this concept.

Your topic is clearly of interest to NHESS and a broad international readership, but the way you present your research needs very thorough attention. I suggest restructuring your manuscript, better outlining your methods and assumptions, adding a dedicated discussion section, and carefully revisiting your concept of validation. You could also help readers appreciating the novelty and advances of your contributions by more clearly and critically assessing what you have achieved here.

__General Remarks__

–Your abstract could do with more detail on how you validated your predictions, and whether they are reliable enough to allow useful flash-flood predictions (or forecasts). What is the eventual output of your prediction and where can this be used in practice?

–The introduction provides some clues why forecasting flash floods is important, but misses opportunities to briefly explain those concepts (especially 'FFG') relevant to your research. Consider making a better case by illuminating more recent case studies of flash floods in South Korea. What is mostly needed for their prediction and why? In this regard, you close the introduction with a somewhat contradictory comment on the need for measuring (antecedent?) soil moisture. Please reconcile that statement and offer a clear overview of your objectives. Which research question is it that you wish to address? Which tools do you use and why?

–The methods section I found difficult to follow. You start of with QPC computation and briefly mention the concept of 'virtual rainfall'. Please elaborate more on that so that readers can reproduce the full stream of your methods. Provide (more) mathematical formulations where appropriate, and please do explain all parameters used (some are not referred to). Why use bankfull discharge? Is that the definition for the minimum discharge to cause flash flooding? Assuming steady, uniform flow may also be problematic for flash floods, and you might want to pick that up in the discussion. Equation

Printer-friendly version

Discussion paper

3 shows a soil-water content balance that you adapt from the SURR model; how well can you constrain each of the five terms? For example, will evapotranspiration as a function of time be relevant for forecasting flash floods? Clearly you want to specify the timescales that you base your forecasts on. I was unsure about the output of your model. Your use of a receiver-operating-characteristic curve indicates that you classify something, but what exactly, remains vague. Please explain in more detail how you labeled the classes of observed events and how you predicted new classes using SURR.

–I suggest changing the order of the methods and study area section. Providing first a general background on the region of interest and the data available before dealing with the method makes more logical sense to me.

–The results section starts off with more methods, uncomfortably emphasizing even more the logical disruption between the early sections of your manuscript. You offer some hydraulic geometry that you derive from a multiple regression model, in which the predictors are clearly correlated. This will need some more robust statistical treatment. Further down the section you mention that the predicted and observed timing of flash floods seem to be roughly similar. This is the first explicit mention of comparing predictions with observed data, and thus the motivation for using ROC curves, I presume. If so, please make sure that this core message comes across much earlier. Again, the time steps or measurement/simulation intervals here are critical. Please elaborate. I am a bit suspicious about Fig. 10. Does basin area somehow play a role in estimating rainfall intensity in any of your models? Finally, your validation (section 4.4) needs to be more convincing. You mention that you tested your method on four observed flash floods between 2005 and 2009. How many cases did you use for training your classifier? Can you show some ROC curves (or other performance metrics) for the testing cases?

–Your study could use a formal discussion section, in which you objectively discuss your methods in the light of their assumptions, limitations, and benefits (or advances)

compared to previous work. Consider reflecting on how accurately SURR produces the necessary input data; how your classification would change for different time intervals; how your classification deals in general with rare events (for which ROC curves might not be the best of performance metrics); and what you consider as possible future improvements to your model.

–Your conclusions mostly summarize your data. You report a high prediction potential, which is partly based on finding the optimal ROC scores in the first place, right? You state that 'The flash flood warning threshold can be best represented as a function of sub-basin area' (page 8/line 27). What does that mean and what is its practical relevance for warning? You may want to report statistical uncertainties for your generalized precipitation-area in this context.

–The reference list appears a bit short. I imagine that other groups must be working on prediction of flash floods elsewhere.

–Figures: #1 is OK, if you add some explanatory detail to the caption; please explain all abbreviations. #2 needs geographic coordinates and larger fonts. #3a and # 4a need units for 'sub-basin area'; are #3b and #4b really necessary? Histogram bins in #5b may be too wide: what is it that you wish to state here? #6 needs explanations of color codes. #7 needs larger fonts and explanation of abbreviations. #8: it is unclear what the minimum and maximum numbers refer to. #9: please explain orange shades. #10: please explain red and blue circles. Overall, you may want to use your captions for informing readers more about the contents and messages of your figures.

–Is Table 1 necessary?

–Please ask a native speaker to check your manuscript. I have noticed numerous formal and potentially ambiguous errors in the text, but these errors are too many to list in detail below. Therefore I only a give only a few examples in the line-specific suggestions below.

__Specific Suggestions (page/line)__

1/8: Delete 'with short duration'. The term 'flash flood' implicates that.

1/9: 'required to cause minor flooding' - Why minor flooding? Please provide a brief definition of what you mean by 'minor' here.

1/12: Please spell out 'ROC'.

1/15: 'highly' should read 'more'?

1/16: 'obtained for rainfall rates of 42, 32 and 20 mm/h' - It is unclear why or how you picked those rates. Please explain.

1/17: 'actual' means 'observed' or 'measured'? Please summarize briefly the results from your validation.

1/20: 'the short-duration flash flood frequently occurred' should read 'the flash floods occur frequently'.

1/24: 'managing flash flood control' - What do you mean by that specifically?

1/25: 'the climate change has increased' could read 'climate change may have likely increased'.

1/27: What sort of 'technology' do you mean? Or did you mean 'methodology' instead?

1/28: 'For deciding flash flood occurrence,' - Unclear.

2/1: 'flash flood vulnerability' - This refers to potential damage. Is that what you meant?

2/5: 'simulation to establish the observed frequency distribution' - Contradictory. Why simulate something to establish observations? Perhaps change the wording here?

2/6: 'comparing forecast flow with flooding flow' - How about 'comparing forecast with observed flows'?

2/7: Delete 'eminent'.

2/9: 'understood by the general public' - It may be useful to briefly explain the concept here.

2/13-15: So what did those studies find out?

2/20: 'the hourly maximum rainfall exceeded 50mm/hr and 60 mm/hr in 2006 and 2011' - Difficult to assess the relevance of these rates without any background information on rainfall characteristics in the region.

2/23: Delete 'exquisitely'. Please also check grammar in this sentence. I think I know what you mean here, but you would be really well advised to seek the help of a native speaker for rephrasing many similar statements in your manuscript.

3/9: What are 'ROC scores'?

3/15: 'method used to compute FFG is the opposite' - So what is the main output of FFG?

3/19: 'over a given duration tr required to' - Please use italics for all parameters that you introduce.

3/24: What is the unit of the 'unit hydrograph peak', if you use differing metric systems? Please attend to Equation 1: in my copy of the PDF it looks as if A is an exponent in the denominator.

4/4: 'which represents current soil conditions' - What do you mean by 'current'? During or before the flash flood?

4/9: 'this model uses estimates soil moisture' - Ambiguous. Does the model use estimates of soil moisture or does it estimate soil moisture itself? That is a big difference.

4/27: Please explain parameters in Equations 4 and 5.

5/5: 'line segments that coincide with the left boundary and upper boundary of the ROC diagram' - You could simply say that, for a perfect prediction, the ROC curve has

to pass through (0, 1) or the upper left point of the graph.

5/7: 'ROC curves associated with real forecasts generally fall between these two extremes and plot above and to the left of the 45-degree diagonal' - Not sure what you mean by a 'real' forecast.

5/15: 'were delineated' - How did you delineate those basins? Their size spans three orders of magnitude, so what was the underlying rationale?

5/19: 'omitted from further analysis' - So you did not consider all basins with reservoirs further?

5/21: 'filtering' - This means you had some preconception about basin area influencing flash-flood potential? It might be good to give more detail here.

5/28: 'soil moisture conditions were estimated' - Please be more specific about the spatial resolution, time intervals, and accuracies of those estimates.

5/30: 'flood information was obtained through different sources, including print and electronic media' - How homogeneous and reliable is that information?

6/1: 'multiple flash flood events' - Perhaps this is something you may wish to elaborate on a bit more?

6/15: 'were investigated and included in the regression equation' - Please describe this in more detail. You note that some of the predictors in your regression model are correlated, but you do not seem to do anything about this.

6/25: 'Threshold runoff values were computed' - How?

6/27: Can you measure runoff rate to one tenth of a mm/h?

6/30: 'flooding season, i.e., July, August and September' - You could explain more about this flooding season in the study area descriptions; international readers might welcome this information.

7/16: 'times of flash flood occurrence computed from the FFG model exhibited satisfactory agreement' - Is it the timing that you wish to classify correctly?

7/21: 'As expected, the minimum ROC score was 0.50' - You can sometimes get lower values than that.

8/12: 'estimated values of 1-hr QFFC' - Do you have measured values for a validation?

8/29: 'optimum threshold for flash flood warning in a sub-basin' - Slight repetition.

9/4: 'which is divided with short and long-duration' - And how do set the threshold between 'short' and 'long'?

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-213, 2017.