NHESS

Ref: NHESS-2022-297

Title: Numerical model derived intensity-duration thresholds for early warning of rainfall-induced debris flows in the Himalayas

Response to Referee #1

Ref.	Comment	Reply
1	I read your manuscript and I appreciated the idea and motivation, but I was left with many questions.	Thank you for your careful consideration and thorough evaluation of our manuscript. We sincerely appreciate your encouragement towards the idea and motivation of this article. Please see the detailed responses below to each of your comments.
2	I do not have a problem with the use of WRF model data to predict hourly rainfall but the prediction needs to be validated with some field data. I understand no hourly data are available in the catchment under study, but are there any data from adjacent areas to verify a degree of correlation? This verification should be done not only for the specific event but in general (e.g., over a whole year) to prove that your approach can be extended and used as a prediction tool. Consider, for example, if your WRF predictions are systematically an overestimation. You still would capture the debris flow events you were seeking, but you would also launch many false alarms.	Thank you for your commendable suggestions. We agree with your concern here. Thanks, we do not have any ground-based rainfall measurement in hourly timestep to validate the WRF outputs. However, considering your suggestion here, we validate the cumulative daily rainfall of WRF outputs with available ground-based precipitation datasets from the India Meteorological Department (IMD). In addition, we validate hourly rainfall from the WRF model with spatially and temporally satellite-derived precipitation data. We also agree to extend the verification over a year (possibly). We agree to the point that WRF predictions may be systematically an overestimation. Of course, this will launch many false alarms of debris flows. However, since using WRF models for landslide early warning is new to the study area, a conservative approach will help first. We would have to tune the model to estimate the rainfall forecast properly.
3	Further, I understand you validated the approach using a rainfall event during which a disaster was actually triggered. This is ok but it is only half of the validation, namely a true positive identification in space and time. What about another event with similar characteristics that did not trigger debris flows in that catchment? Or the same event but in an adjacent catchment where no debris flows occurred? To be	Thank you for this very comprehensive and helpful comment. During the 2013 North India Floods, more than 6000 landslides were triggered in Uttarakhand, India. Though it is difficult to identify a catchment where no debris flows occurred due to this extreme rainfall event, we try our best to identify

	usable as a warning system, your approach should also be able to identify true negatives in space and time.	a catchment with fewer debris flows and perform the simulation.
4	Further, you had to make assumptions due to lack of field data (e.g., on the pre-rainfall moisture) but you did not discuss how reasonable your choice was or how sensitive the result is to a change in the chosen value. In other words, where does the 5% moisture come from? Is it supported by field or lab experiments? What changes if you use a moisture of 0% or 20%?	Thank you for your careful examination and constructive suggestions. We agree with your concern here. You are correct in pointing out that we had to make assumptions due to a lack of field data. Since the initial conditions could be sensitive to the triggering time of debris flows, we had to decide that carefully. We run a decadal simulation of rainfall-runoff-infiltration using daily timesteps of rainfall (data from IMD) from 2003 to 2013. We used the initial moisture content from the results. Thanks for the thoughts. We also perform additional numerical simulations to test the sensitivity of initial moisture content and include these results in the revision stage.
5	Finally, the DEM resolution. 30 m does not really seem great at your scale, with a catchment of few km. I agree that resampling cannot improve the result (because a smooth DEM remains smooth after resampling), but what about an actual high-resolution DEM that more closely follows the roughness of the morphology? If not available in this catchment, couldn't the authors study this sensitivity in another location with better data, to assure the reader that the result remains reasonable?	Thank you for this very useful suggestion. We run a parameter sensitivity analysis using different resolutions of DEM to assure the reader that the result remains reasonable.