

## Replies to Reviewer 2

I've read and carefully evaluated the manuscript and my opinion is that it needs major revisions before being published in HESS. Although the scientific significance is high, I think that the scientific quality is affected by several shortcomings. The manuscript is generally well written, the presentation quality is fair and could be improved in some parts. Please find hereafter my main concerns, divided among general and specific issues.

### GENERAL ISSUES

1- The authors should clearly differentiate this work from their recently published "Evaluation of Remotely Sensed Soil Moisture for Landslide Hazard (IEEE-JSTARS 2019)" and avoid that some introductory parts read very similar.

Reply: Agreed. As the reply to Reviewer 1, in the introduction, some parts will be removed to reduce the similarity from the previous paper on explaining the existing research gaps. The shortcomings of the previous study will be added in the updated manuscript, to clearly explain the necessity and novelty of this study (i.e., the need of using high spatial (both horizontally and vertically) and temporal resolution soil moisture products for landslide application).

2- I am somehow concerned about the dataset used. Emilia Romagna is one of the Italian regions with the best environmental datasets, some of them also publicly available for free. Therefore, I wonder why a dataset period was chosen in which only a soil moisture station is available, and why a manuscript submitted in 2019 relies on datasets until only 2015 (all dataset used extend almost until present days). DEMs of the region are available at finer spatial resolutions (10m and 20m): why using a 90m resolution SRTM DEM? Lastly, the landslide dataset seems largely incomplete. Many works on the same test site (see following comment) used larger landslide dataset for the same time period. This shortcoming may let the readers question about the significance of results obtained.

Reply: The reason for choosing the 10-year period of 2006-2015 is because although there are 19 soil moisture sensors installed in the region, only the San Pietro Capofiume station can provide the longest continuous valid data (2006 to early 2017). We have checked all the other stations, and they are either absent from valid data (e.g., have very big data gaps) or do not have data at all. This will be further clarified in the updated version of the manuscript.

The reason for using the SRTM DEM data is because it covers most part of the world and is free to download, which makes the study applicable globally.

The landslide data is collected from the Emilia Romagna Geological Survey. There is a total of around 800 events occurred during the 10-year study period. The reason for choosing only one-fifth of the events (i.e., 157 events) is because we have to filter the data based on: 1) rainfall-induced only; 2) records with high spatial-temporal accuracy (exact date and coordinates); and 3) landslide only (other events such as debris flow, rockfall and collapse are excluded).

3- The references of the manuscript are very biased. I think the authors cited almost their whole scientific production (e.g. at line 294: are three references from the same author necessary?), while they almost ignored what has been published on the same subject and in the same case of study. For instance: landslide characteristics could be better described making reference to

some previous works (e.g. Bertolini et al., 2005; Rossi et al., 2010). Regional scale rainfall thresholds for the Emilia Romagna have been already published by Berti et al. (2012) using an I-D approach and by Martelloni et al. (2012) using antecedent rainfall. Regional scale landslide warning systems for the Emilia Romagna Region have been addressed in several papers (e.g. Lagomarsino et al., 2013; Segoni et al. 2018a). Lagomarsino et al. (2015) compared an I-D threshold model and an antecedent rainfall threshold model concluding that in Emilia Romagna the latter provides better performances, probably due to the complex hydrologic response of the hillslopes after rainfalls. Segoni et al., 2018b (already in your reference list) tested that the performances of the Emilia Romagna threshold system could be improved by integrating basin-scale soil moisture estimated by means of TOPKAPI model. I think all those antecedent works could be used to properly “set the stage” for your research. Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *Journal of Geophysical Research: Earth Surface*, 117(F4). Bertolini, G., Guida, M. & Pizziolo, M. *Landslides* (2005) 2: 302. <https://doi.org/10.1007/s10346-005-0020-1> Segoni, S., Rosi, A., Fanti, R., Gallucci, A., Monni, A., & Casagli, N. (2018). A Regional-Scale Landslide Warning System Based on 20 Years of Operational Experience. *Water*, 10(10), 1297. Lagomarsino, D., Segoni, S., Fanti, R., & Catani, F. (2013). Updating and tuning a regional-scale landslide early warning system. *Landslides*, 10(1), 91-97. Martelloni, G., Segoni, S., Fanti, R., & Catani, F. (2012). Rainfall thresholds for the forecasting of landslide occurrence at regionalscale. *Landslides*, 9(4), 485-495. Lagomarsino, D., Segoni, S., Rosi, A., Rossi, G., Battistini, A., Catani, F., & Casagli, N. (2015). Quantitative comparison between two different methodologies to define rainfall thresholds for landslide forecasting. *Natural Hazards and Earth System Sciences*, 15(10), 2413-2423. Rossi, M., Witt, A., Guzzetti, F., Malamud, B. D., & Peruccacci, S. (2010). Analysis of historical landslide time series in the Emilia-Romagna region, northern Italy. *Earth Surface Processes and Landforms*, 35(10), 1123-1137.

**Reply:** We thank the reviewer for the comments. We will modify the references in the updated manuscript, so that a wide range of relevant papers are cited.

4- At this scale of analysis, the attempt to relate modeled soil moisture to a single instrumented site is a too big stretch in my opinion. Please, consider also that the sensor is located in a completely different setting (wide alluvial plain) than the territory typically affected by landslides (hills and mountains). I think the trends of soil moisture could be largely unrelated (as also the authors stated at line 61) and the example at line 328 (500km radius) would not hold in a case study characterized by many differences and peculiarities like Emilia Romagna. The authors could maybe cite other authors that attempted to establish empirical correlations of hydrological variables in Emilia Romagna (Segoni et al., 2018 with soil moisture; Martelloni et al., 2013, with snowpack thickness), however they calibrated the relationships over smaller territorial units, not over the whole region. I think these works could be used to partially defend the approach used in the manuscript, but I don't think they can completely clear the feeling that just one single instrument for the whole region is insufficient.

**Reply:** Agreed. As suggested by Reviewer 1 (“it is well known that soil moisture is strictly related to rainfall, and I was wondering how accurate are the WRF simulated rainfall? A comparison between observed and simulated rainfall should be carried out to have a better understanding of the quality of WRF model in the study area”). We will add the rainfall

evaluation work in the updated manuscript, which will provide useful indicators for the accuracy (both spatially and temporally) of the WRF estimated soil moisture in the study region.

5- The authors should be very careful in providing unbiased, objective and humble points of view. The feeling is that in some parts of the manuscript they are overreaching when describing the results obtained (e.g. “outstanding” at line 479). Indeed, in my opinion the results are questionable. Beside the issue of using an instrument located in the alluvial plain to model landslide occurrence in very different climatic, hydrologic and geo-morphologic settings, there is a clear problem of result evaluation: most of the results are presented as graphics where it is difficult to ascertain the goodness of the model fitting because a long dataset is compressed in a small figure and also a qualitative evaluation is hard (sections 4.1. and 4.2). Some quantitative validation is mandatory to better evaluate the results. We need to know the differences, how big they are, where/when are located and why they are present. Also, about abstract, results and discussion: I don't think WRF modelled soil moisture has been properly evaluated for landslide monitoring purposes (line 464-465). This work in my opinion can be considered a preliminary attempt towards that direction, but to reach the goal more and better data are needed, together with a more thorough and quantitative evaluation of the results. I suggest that the authors rephrase their statements.

Reply: Agreed. The ‘too optimistic’ sentences will be rephrased throughout the paper.

## **SPECIFIC ISSUES**

18. “landslide threshold model” is a very generic term. Please, be more specific.

Reply: Agreed. This will be modified.

40-42. Please adjust this sentence and provide more references if necessary. Caine was the first to establish an I-D threshold, but to my knowledge that threshold has never been used for a warning system. In addition, national scale landslide warning systems are not so common and not so many examples of prototypal or operation applications exist in the literature (e.g. Krogly et al., 2018; Rosi et al. 2016; Auflic et al., 2016). Indeed, threshold-based landslide warning systems are usually established for smaller areas (e.g. basins or regions or small alert zones), see e.g. Devoli et al. (2018), Baum and Godt (2010), Mathew et al. (2014) : : :

Reply: Agreed. The sentence will be adjusted.

149-150. “weak earth units” is unclear. Please, rephrase.

Reply: Agreed. This will be rephrased.

237 “an improved”

Reply: This will be corrected.

279-280. Not clear, please rephrase.

Reply: This will be rephrased.

303-304. I think the concepts of TP/TN/FP/FN are quite established, no need to make reference to other works.

Reply: Agreed. The reference will be removed.

313. Maybe I'm mistaken, but I don't think at this point the slope degree groups have been presented yet.

Reply: Agreed. The sentence will be modified.

342. Please rephrase: "very well" cannot be used (see also general comments).

Reply: Agreed. This will be rephrased.

356-359. So, you are saying that the dataset has a bad quality? Maybe the dataset needs to be smoothed?

Reply: We suspect it's due to the sensing failure at the deep-layer. As suggested by Reviewer 1, we will further clarify this issue.

370. Please, revise the English.

Reply: This will be revised.

378. Actually, in my opinion you don't have a reliable benchmark, so you can only say that the models provided different results, you cannot say which one provided the best results.

Reply: Agreed. This will be rephrased.

387. This part is very important, but you did not introduce it appropriately. This is strange, because in the introduction you cited many relevant papers (Glade et al., 2000; Brocca et al., 2008; Segoni et al., 2018; Bogaard and Greco 2018). Maybe you should spend a few words saying that previous works demonstrated that in complex geomorphologic settings (as Emilia Romagna) a rainfall threshold approach is too simple and more hydrologically-driven approaches need to be established.

Reply: Agreed. This sentence will be expanded as suggested by the reviewer.

396. I disagree. The relationship between the slope angle and the landslide triggering is not so straightforward and it depends on the landslide typology. E.g. many slow earth flows (which are abundant in Emilia Romagna) can occur also on very low slope gradients.

Reply: Agreed. This sentence will be rephrased.

407. This choice is questionable. The landslide susceptibility does not necessarily have to be equally represented in the territory.

Reply: There are different ways to group the slopes. In this study, three groups have been defined with similar sizes so that relatively reliable results could be achieved from the statistical point of view.

422-426. This is quite obvious, I wouldn't devote so much space to this.

Reply: Agreed. This part will be shortened.

427. Why are you mentioning shallow landslides? Until now, I figured out that you are trying to model landslides in general.

Reply: Agreed. This will be modified.

430-431. Please, rephrase this sentence.

Reply: Agreed. This will be rephrased.

453-454. Theoretically, the conditions at the sliding surface should be the ones with the best performances. Therefore, either you have very shallow landslides, or your results are not good at assessing the real soil moisture conditions that are actually triggering/predisposing landslide initiation.

Reply: One potential reason is that, the conditions at the sliding surface are important, but the soil moisture above it is also important (the loading should be heavier with more water in the upper layer soil). We will add more discussion about the results.

475-476. Please rephrase.

Reply: This will be rephrased.

479. Do not use “outstanding”.

Reply: Agreed. This will be modified.

481. I failed to quickly check the correlation coefficient. Please, clearly write where it can be found by the readers.

Reply: Agreed. The correlation coefficient information will be added here.

499. To my experience, the applicability of a similar system in a warning system would be very limited because it would have a poor spatial resolution: warning would be issued over the whole region, thus having limited actual applicability.

Reply: Agreed. The WRF modelled soil moisture data could provide useful information for landslide studies, but for early warning system, a lot of additional information (e.g., particularly high resolution data) will be needed. The sentence will be rephrased.

628. Please, check.

JEMEC AUFLIČ M, ŠINIGOJ J, KRIVIC M, PODBOJ M, PETERNEL T, KOMAC M, 2016. Landslide prediction system for rainfall induced landslides in Slovenia (Masprem). GEOLOGIJA 59/2, 259-271 Baum RL, Godt JW (2010). Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides, 7(3):259– 272. Krøgli, I. K., Devoli, G., Colleuille, H., Boje, S., Sund, M., and Engen, I. K.: The Norwegian forecasting and warning service for rainfall- and snowmelt-induced landslides, Nat. Hazards Earth Syst. Sci., 18, 1427-1450, <https://doi.org/10.5194/nhess-18-1427-2018>, 2018 Martelloni, G., Segoni, S., Lagomarsino, D., Fanti, R., & Catani, F. (2013). Snow accumulation/melting model (SAMM) for integrated use in regional scale landslide early warning systems. Hydrology and Earth System Sciences, 17(3), 1229- 1240. Mathew, J., Babu, D. G., Kundu, S., Kumar, K. V., & Pant, C. C. (2014). Integrating intensity–duration-based rainfall threshold and antecedent rainfall-based probability estimate towards generating early warning for rainfall-induced landslides in parts of the Garhwal Himalaya, India. Landslides, 11(4), 575-588. Rosi, A., Peternel, T., Jemec-Auflič, M., Komac, M., Segoni, S., & Casagli, N. (2016). Rainfall thresholds for rainfall-induced landslides in Slovenia. Landslides, 13(6), 1571-1577.

Reply: The references will be checked.