Replies to comments from reviewer #1

We thank the Reviewer #2 for his/her kind comments.

Line 9: Remote sensing is not a tool. Please rephrase the sentence

R: We rephrased with: "*In these areas, remote sensing can represent an important detection and monitoring process to predict landslide occurrence...* ".

Line 10: Is it forecasting? It seems like prediction. There is difference between prediction, forecast and hindcasting. Please be clear what is actually being done

R: You're perfectly right. Indeed, here we perform a prediction of rainfall-induced landslides.

Line 15: Any specific reason to use only these products? Why not check other climatic reanalysis datasets?

R: We decided to test these two products as we wanted to highlight the added value of using satellite data. The choice of GPM and SM2RAIN-derived rainfall lies in the fact that we wanted to test the capabilities of two different retrieval approaches, i.e. a state-of-the-art classical satellite rainfall product and a novel technique that estimates rainfall through soil moisture observations, to provide useful information for landslide applications. Moreover, we decided to not consider other climatic reanalysis data as we wanted to highlight the feasibility of using rainfall estimates in available in near-real time for operational purposes. Indeed, GPM data are provided within 4 hours after sensing (the Early run product here considered), a near-real time product based on SM2RAIN is currently under development within HSAF, while ERA 5 for instance is provided 5 days after the model run.
We better clarify this point at lines 188-191:
"*The satellite rainfall products here considered allow to test the capabilities of two different retrieval algorithms, i.e. a state-of-the-art classical satellite rainfall product and a novel technique that estimates rainfall through soil moisture observations. Other reanalysis data were not considered as we wanted to investigate the use of near-real time satellite products for a possible future operational application.*"

Line 19: Most of the landslide data in this region is on a daily basis. How did you find the hourly based landslide location. Was the historical landslide events collected based on high temporal resolution satellite images?

R: We didn't use satellite images to locate the landslides. We collected landslide information from online newspapers and magazines (ON), blogs (BG), technical reports (TR) made available by the Geological Survey of India (GSI, www.gsi.gov.in), scientific journals (SJ) and social media (SM). For the majority (about 64%) of the landslides we found the occurrence time or the part of the day.

Line 34: Too many references, for one generic statement. Either mention the country and the corresponding citation or mention only the notable review papers.

R: We removed the following references: Ruiz-Villanueva et al. 2011, Rosi et al., 2016, Peruccacci et al. (2017), but we prefer to leave the others.

Line 59: You can also mention the recent work of Dikshit et al. (2020) where they argued the necessity to use satellite based products in the Indian Himalayan region.
Dikshit, A. et al. Rainfall Induced Landslide Studies in Indian Himalayan Region: A Critical Review. Appl. Sci. 2020, 10, 2466.

R: We added: "*In a recent work, Dikshit et al. (2020b) argued the necessity to use satellite-based products in the Indian Himalayan region*."

Line 76: See, here the word "predict" is used. Please be thorough in the entire manuscript.

R: Done.

Line 76: English problem.

R: We have modified the sentence as follows: "*we collected information on rainfall-induced landslides in a new catalogue*."

Line 99: You can also add the Koppen climate classification.

R: It could be interesting to use the Koppen classification to investigate the role of the climate. In India there are at least 5 Koppen climate classes, but the limited number of landslides available for this work prevents such analysis.

Line 101: What do you mean by statistically significant? Do you mean total number of landslides in all of India or total number in a specific region? If so, how many and at what area?

R: Here, "statistically significant" means larger than a minimum number of landslides to define reliable regional thresholds as indicated in Peruccacci et al. (2012). In that work we found that this number is between 75 and 100.

Line 111: Can you add a list of landslide events included and discarded in the analysis?

R: Definitely, we can share the catalogue, but the discarded events were not included in it, since they were not useful for our purposes. Events that did not have the required spatial and temporal accuracy were not stored.

Line 118: I don't think Rajasthan has a problem of landslide occurrences. Kindly check.

R: Unfortunately, this unique landslide killed one and injured six residents in the foothill zone of Amagarh Hill.

Line 121: Could you elaborate on this?

R: We changed "*social incidence*" in "*impact on the population*".

Line 122: I assume that this finding was based after 50% of the landslide data was discarded. If they were not discarded, does it depict the same results?

R: As we stated in a previous answer, the discarded events were not stored and therefore were not available for this analysis.

Line 125: Please check Uttarakhand landslides should be a big contributor on fatalities

R: What is shown in Fig. 2b is relative to our catalogue, which was compiled with the criteria described above. In Uttarakhand, we had many problems in finding the landslide coordinates, since many of the failures occurred "along" long routes, and were consequently discarded from the catalogue.

Line 128: Were these reliable, in terms of location and dates. Did you cross-check it with newspapers?

R: As per fig. 3a, social media accounted only for 0.5%, and they were checked with newspapers. Social media are especially useful to find details on the time of occurrence of landslides.

Line 135: Does this sort of classification have an effect on the model outcomes or is it providing an analysis on the landslide type in the region

R: Due to the limited number of landslides in each class, the classification is only providing an analysis on the landslide type.

Line 150: Reason for using this interpolation? Why not krigging?

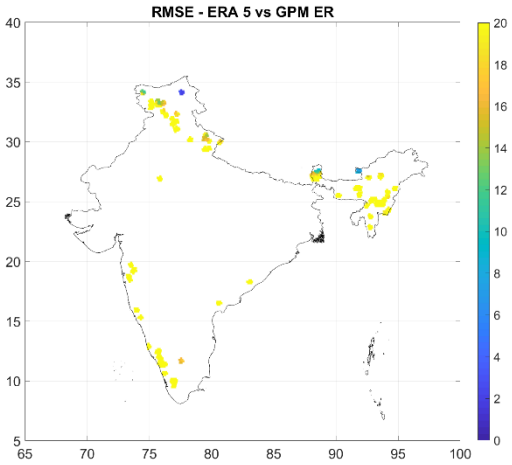
R: We used the IMD product as it is provided by the Indian Meteorological Department, we did not perform any other processing step to ground data.

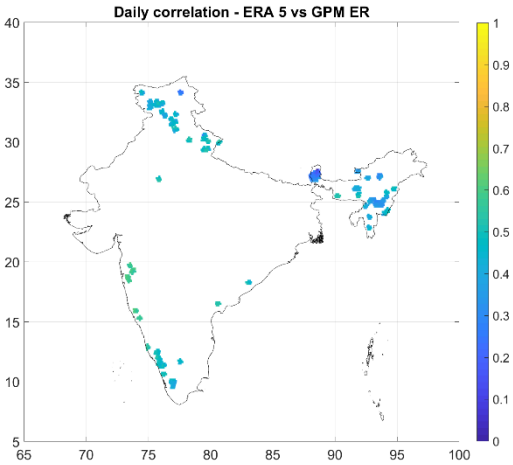
Line 156: for which region was these studies conducted?

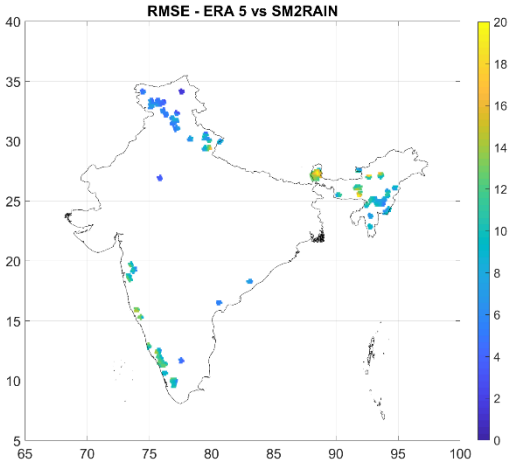
R: Europe.

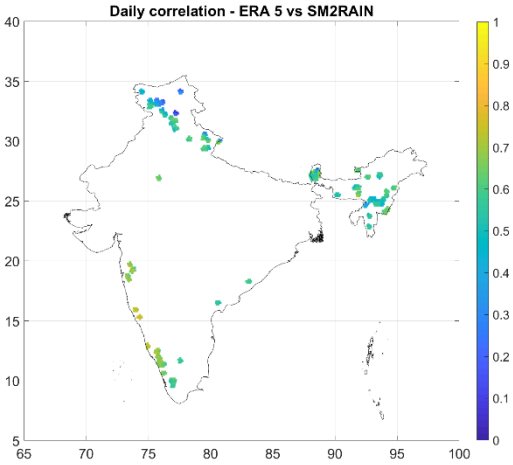
Line 183: Here, the integration and the way rainfall is estimated is not clear. Although, you provide references that these data and processes seem to work, was it effective for determining rainfall in Indian context and if they were what were the uncertainities in different parts of the country.
This brings to my previous comment, why not use climatic reanalysis dataset like ERA5, by that way you could have processed more data points?

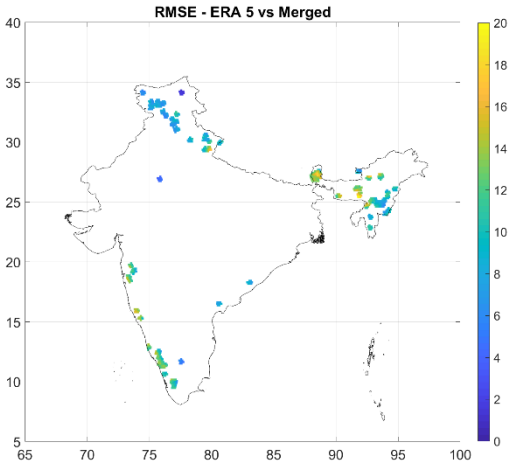
R: In order to highlight the fruitful effect of the integration between the two datasets, we performed a very simple test, estimating the daily correlation coefficient and RMSE, using ERA5 as benchmark. As a result, the performance of the integrated product improved both in terms of correlation and RMSE on a daily scale, throughout the study area. The results can be observed in the following figures.

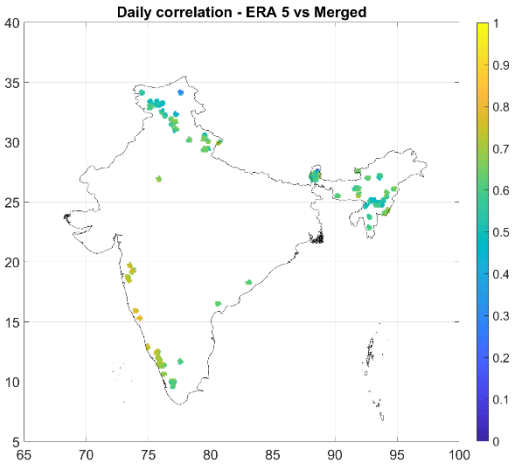












For what concerns ERA 5, please refer to our previous comment.

Line 201: The buffer of 20km was it used for entire country or was it used in specific parts of the country? Why not put a varying buffer size dependent on the rain gauge density.

R: We applied the same buffer radius all over the country. For comparison with gridded satellite data, we used gridded ground rainfall data provided by IMD, which doesn't contain any information on the rain gauge density.

Line 250: English problem.

R: We rephrased it as follows: "*These two products exhibit the highest number of rainy days along the whole year*"

Line 277: After reading this section, it looks that only the rainfall parameter is being predicted, but not the locations? Thus, the title and the wording in the manuscript should change to "prediction of rainfall thresholds for landslide occurrences" rather than "forecasting of landslides"

R: We agree, and we changed the title and the wording in the manuscript.

Line 325: Use a milder word. I understand this is an excellent work, but I don't think pioneering is the right word.

R: Done.

Figure 2: Please mention the years.

R: Done.

Figure 7: If you could use different colors for better clarity, that would be appreciated.

R: We prefer to keep these colors.